



MontCAS

Criterion-Referenced Test

(MONTANA CRT)

2010–11

Technical Report

TABLE OF CONTENTS

CHAPTER 1.	OVERVIEW OF THE MONTANA CRITERION-REFERENCED TEST.....	1
1.1	Purpose of the Assessment System.....	1
1.2	Purpose of this Report	2
CHAPTER 2.	ASSESSMENT AND TEST DEVELOPMENT PROCESS	3
2.1	Test Specifications.....	3
2.1.1	Criterion-Referenced Test.....	3
2.1.2	Item Types	3
2.1.3	Description of Test Design	3
2.2	Reading Test Specifications	4
2.2.1	Standards.....	4
2.2.2	Item Types	4
2.2.3	Test Design	4
2.2.4	Blueprints (Distribution of Points Across Standards)	5
2.2.5	Depth of Knowledge	5
2.2.6	Passage Types	6
2.3	Mathematics Test Specifications	8
2.3.1	Standards.....	8
2.3.2	Item Types	8
2.3.3	Test Design	8
2.3.4	Blueprints (Distribution of Points Across Standards)	9
2.3.5	Depth of Knowledge	9
2.3.6	Use of Calculators.....	10
2.4	Science Test Specifications	11
2.4.1	Standards.....	11
2.4.2	Item Types	11
2.4.3	Test Design	12
2.4.4	Blueprints (Distribution of Points Across Standards)	12
2.4.5	Depth of Knowledge	12
2.4.6	Use of Calculators and Reference Sheets.....	13
2.5	Test Development Process.....	13
2.5.1	Item Development.....	13
2.5.2	Item Reviews at Measured Progress	15
2.5.3	Item Reviews at State Level.....	15
2.5.4	Bias and Sensitivity Review.....	15
2.5.5	Reviewing and Refining.....	16
2.5.6	Item Editing	16
2.5.7	Item Selection and Operational Test Assembly	16
2.5.8	Operational Test Draft Review	17
2.5.9	Alternative Presentations	17
2.6	Test Sessions	18
CHAPTER 3.	TEST ADMINISTRATION	19
3.1	Responsibility for Administration	19
3.2	Administration Procedures	19
3.3	Participation Requirements and Documentation	20
3.3.1	Students with Disabilities.....	21
3.4	Administrator Training	21
3.5	Documentation of Accommodations	21
3.6	Test Security and Administration Irregularities.....	22
3.7	Test Administration Window	23
3.8	Service Center	23
CHAPTER 4.	SCORING	24
4.1	Machine-Scored Items	24
4.2	Person-Scored Items	25
4.2.1	Scoring Location and Staff	26
4.2.2	Reader Recruitment and Qualifications	28
4.2.3	Methodology for Scoring Polytomous Items	29

4.2.4	Reader Training	30
4.2.5	Leadership Training	33
4.2.6	Monitoring of Scoring Quality Control.....	34
CHAPTER 5.	CLASSICAL ITEM ANALYSIS.....	39
5.1	Classical Difficulty and Discrimination Indices.....	39
5.2	Differential Item Functioning	42
5.3	Dimensionality Analysis.....	43
CHAPTER 6.	IRT SCALING AND EQUATING.....	47
6.1	Item Response Theory	48
6.2	Item Response Theory Results	49
6.3	Equating.....	51
6.4	Equating Results.....	51
6.5	Achievement Standards	52
6.5.1	Distributions.....	52
6.6	Scaled Scores.....	52
6.6.1	Description of Scale	52
6.6.2	Calculations.....	53
6.6.3	Distributions.....	54
CHAPTER 7.	RELIABILITY	55
7.1	Reliability and Standard Errors of Measurement.....	56
7.2	2010–11 Subgroup Reliability	57
7.3	Reporting Subcategory Reliability	57
7.4	Interrater Consistency	57
7.5	Reliability of Performance Level Categorization	58
7.5.1	Decision Accuracy and Consistency Results	60
CHAPTER 8.	SCORE REPORTING.....	62
8.1	Decision Rules.....	63
8.2	Static Reports.....	63
8.2.1	Student Report	63
8.2.2	Summary Reports.....	63
8.3	Montana Analysis and Reporting System (MARS).....	65
8.3.1	User Accounts	65
8.4	Interactive Reports.....	65
8.4.1	Roster Report	66
8.4.2	Performance Level Summary.....	67
8.4.3	Item Analysis Data.....	67
8.4.4	Longitudinal Data Report.....	67
8.5	Interpretive Materials and Workshops.....	67
8.6	Quality Assurance	67
CHAPTER 9.	VALIDITY.....	69
REFERENCES		71
APPENDICES.....		73

Appendix A	Analysis and Reporting Decision Rules
Appendix B	Participation Rates
Appendix C	Accommodation Frequencies
Appendix D	Item Review Committee Members
Appendix E	Item Level Classical Statistics
Appendix F	Item Level Score Distributions
Appendix G	Number of Items Classified Into DIF Categories
Appendix H	Item Response Theory Calibration Results
Appendix I	TCCs and TIFs
Appendix J	<i>b</i> -Plots
Appendix K	Analyses of Equating Items (Delta & Rescore Analyses)
Appendix L	Score Distributions
Appendix M	Raw to Scaled Score Look-up Tables

Appendix N	Classical Reliability
Appendix O	Interrater Agreement
Appendix P	Decision Accuracy and Consistency Results
Appendix Q	Report Shells
Appendix R	Sample Interactive Reports

CHAPTER 1. OVERVIEW OF THE MONTANA CRITERION-REFERENCED TEST

1.1 PURPOSE OF THE ASSESSMENT SYSTEM

The Montana Criterion-Referenced Test (CRT) was developed in accordance with the following federal laws: Title 1 of the Elementary and Secondary Education Act (ESEA) of 1994, P. L. 103–382, and the No Child Left Behind Act (NCLB) of 2001.

In the spring of 2011, Montana students in grades 3 through 8 and 10 participated in the MontCAS Criterion-Referenced Test (Montana CRT) in reading, mathematics, and science. The purpose of this assessment is to measure students' achievement as articulated by Montana content standards and grade-level expectations.

All Montana students enrolled in accredited schools are required to participate in either the Montana CRT or the Montana CRT-Alternate. The vast majority of students participate in the CRT using standard administration procedures. However, an array of standard accommodations is available to any student, with or without disabilities, when such accommodations are necessary to allow the student to demonstrate his/her skills and competencies. Standard accommodations are not considered to change the constructs being measured and may be provided to students as necessary for any or all of the reading, mathematics, or science portions of the assessment. Students' tests are scored the same way whether they take the test using standard accommodations or not.

In addition to standard accommodations, other accommodations for the Montana CRT are available to students when specified in their IEPs, 504 plans, or LEP plans. These other accommodations are referred to as nonstandard accommodations. Because they are considered to alter the constructs being measured, they do affect a student's score on the CRT. When a nonstandard accommodation is used, the student's score is reported as the lowest possible for that content area (e.g., a scaled score of 200 will fall into the Novice performance level). Nonstandard accommodations may be provided in reading, mathematics, or science, as dictated by the student's IEP, 504 plan, or LEP plan.

A small percentage of students participate in the statewide assessment program by taking the CRT-Alternate. Students with significant cognitive disabilities, who are working toward alternate academic achievement standards, as documented in their Individualized Education Program (IEP), are eligible to take the CRT-Alternate. Technical characteristics of the CRT-Alternate program are described in a companion technical report.

Montana grade and content area combination CRT instruments are based on and aligned to Montana's content standards, benchmarks, and grade-level expectations in reading, mathematics, and science. Montana educators worked with the Montana Office of Public Instruction (OPI) and Measured Progress to develop test items that assess how well students have met Montana grade-level expectations for each content area. In

addition, Northwest Regional Educational Laboratory (NWREL) performed an independent alignment study for mathematics and reading in 2006 and for science in 2007. NWREL's alignment studies can be found on the OPI's Web site at www.opi.mt.gov/assessment.

Montana CRT scores are intended to be useful indicators of the extent to which students have mastered material outlined in Montana reading, mathematics, and science content standards, benchmarks, and grade-level expectations. Each student's Montana CRT score should be used as part of a body of evidence regarding mastery and should not be used in isolation to make high-stakes decisions. Montana CRT scores are more reliable indicators of program success when aggregated to school, system, or state levels, particularly when monitored over the course of several years.

1.2 PURPOSE OF THIS REPORT

This report describes technical aspects of the Montana CRT in an effort to contribute to the accumulation of validity evidence to support Montana CRT score interpretations. Because the interpretations of test scores, not the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations (American Educational Research Association [AERA], American Psychological Association & National Council on Measurement in Education, 1999). Subsequent chapters of this report discuss test development, test alignment, test administration, scoring, equating, item analyses, reliability, scaled scores, performance levels, and reporting. Each of these topics contributes important information toward establishing the validity of the assessment program. Note, however, that this report does not include certain aspects of a comprehensive validity argument that could also be important to consider when drawing conclusions about validity. (For instance, additional sources of validity evidence might speak to the extent to which Montana CRT scores converge with other measures of the same or similar constructs and diverge from measures of different constructs, or they might examine consequences that arise from scores at the student, school, district, and state levels.)

Historically, some parts of technical reports may have been used by educated laypersons, but the intended audience was experts in psychometrics and educational research. This edition of the Montana CRT Technical Report attempts to make information more accessible to educated laypersons by providing more thorough descriptions of general categories of information. While making some information more accessible, we have also purposely preserved the depth of technical information provided. The reader will find that some discussions and tables continue to require a working knowledge of measurement concepts, such as “reliability” and “validity,” and statistical concepts, such as “correlation” and “central tendency.” To fully understand some of the data presented, the reader will have to possess a basic understanding of advanced topics in measurement and statistics.

CHAPTER 2. ASSESSMENT AND TEST DEVELOPMENT PROCESS

2.1 TEST SPECIFICATIONS

2.1.1 Criterion-Referenced Test

Items on the Montana Criterion-Referenced Test (CRT) are developed specifically for Montana and are directly linked to Montana's content standards. These content standards are the basis for the reporting categories developed for each content area and are used to help guide the development of test items. No other content or process is subject to statewide assessment. An item may address part, all, or several of the benchmarks within a standard.

2.1.2 Item Types

Montana educators and students are familiar with the types of items used in the assessment program. The types of items and their functions are described below:

- **Multiple-choice (MC)** items are used to provide breadth of coverage within a content area. Because they require no more than a minute for most students to answer, multiple-choice items make efficient use of limited testing time and allow for coverage of a wide range of knowledge and skills.
- **Short-answer (SA)** mathematics items are used to assess students' skills and abilities to work with brief, well-structured problems that have one or a very limited number of solutions (e.g., mathematical computations). Short-answer items require approximately two minutes for most students to answer. The advantage of this type of item is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.
- **Constructed-response (CR)** items typically require students to use higher-order thinking skills—evaluation, analysis, summarization, and so on—to construct satisfactory responses. Constructed-response items take most students approximately five to ten minutes to complete. Note that the use of released Montana CRT items to prepare students to respond to constructed-response items is appropriate and encouraged.

2.1.3 Description of Test Design

The Montana CRT is structured using both *common* and *field-test* items. Common items are taken by all students in a given grade level. Student scores are based only on common items. In addition, field-test items are divided among the four forms of the test for each grade level. Each student takes only one form of the test and therefore answers a fraction of the field-test items. Field-test items are not identifiable to test takers and have a negligible impact on testing time. Because all students participate in the field test, it

provides the sample size (750–1500 students per item) needed to produce reliable data that can be used to inform item selection for future tests.

2.2 READING TEST SPECIFICATIONS

2.2.1 Standards

The test specifications/blueprint for reading is based on Montana’s reading content standards, which identify five Montana content standards that apply specifically to reading and reading comprehension. Those content standards are listed below:

- **Reading Standard 1:** Students construct meaning as they comprehend, interpret, and respond to what they read.
- **Reading Standard 2:** Students apply a range of skills and strategies to reading.
- **Reading Standard 3:** Students set goals, and monitor and evaluate their reading progress. (This standard cannot be measured with a traditional paper-pencil test.)
- **Reading Standard 4:** Students select, read, and respond to print and non-print materials for a variety of purposes.
- **Reading Standard 5:** Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audience.

2.2.2 Item Types

The Montana CRT in reading includes a mix of multiple-choice and constructed-response items. Constructed-response items require students to write answers consisting of one or more paragraphs. Each type of item is worth a specific number of points in the student’s total reading score, as shown in Table 2-1.

Table 2-1. 2010–11 Montana CRT: Item Types

<i>Item type</i>	<i>Possible score points</i>
MC	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; CR = constructed-response

2.2.3 Test Design

Table 2-2 shows the numbers of multiple-choice and constructed-response items for grades 3 through 8 and 10.

Table 2-2. 2010–11 Montana CRT: Common Reading Items—Grades 3–8 and 10

Grade	Session 1	Session 2	Session 3	Total	
				MC	CRs
3–8	19 MC, 1 CR	14 MC	19 MC, 1 CR	52	2
10	19 MC, 1 CR	14 MC	19 MC, 1 CR	52	2

MC = multiple-choice; CR = constructed-response

2.2.4 Blueprints (Distribution of Points Across Standards)

Table 2-3 shows the distribution of points across content standards.

**Table 2-3. 2010–11 Montana CRT:
Reading Specifications/Blueprint Grades 3–8 and 10**

<i>Number of points for the common (scored) test:</i>							
52 MC items + 2 CR items = 60 points							
Percent point distribution by content standard*							
Content standards	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Standard 1	34%	34%	34%	34%	34%	34%	25%
Standard 2	30%	30%	30%	30%	30%	30%	32%
Standard 3							
Standard 4	18%	18%	18%	18%	18%	18%	22%
Standard 5	18%	18%	18%	18%	18%	18%	22%
*Because percents are rounded to the nearest whole number, not all sums add to 100%. Note: Standard 3 cannot be measured with a traditional paper-pencil test.							
Target point distribution by content standard (acceptable range)							
Content standards	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Standard 1	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	15 (13–17)
Standard 2	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	19 (17–21)
Standard 3							
Standard 4	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	13 (11–15)
Standard 5	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	13 (11–15)

MC = multiple-choice; CR = constructed-response

Four-point items: Each test contains two four-point constructed-response items. In any given year, the two items will measure two different standards. From year to year, those standards may change.

One-point items: The number of one-point items per content standard will vary from year to year depending on which two standards are measured by the four-point items. (The number of total points per standard falls within the acceptable range from year to year.)

2.2.5 Depth of Knowledge

Each item on the Montana CRT in reading is assigned a depth of knowledge (DOK) level. The depth of knowledge level reflects the complexity of mental processing students use to answer an item. Depth of knowledge is not synonymous with difficulty. Each of the levels is described below.

- **Level 1 (Recall).** This level requires students to receive or recite facts or to use simple skills or abilities. Items require only a shallow understanding of the text presented and often consist of verbatim recall from text, slight paraphrasing of specific details from the text, or simple understanding of a single word or phrase.
- **Level 2 (Skill/Concept).** This level includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Inter-sentence analysis of inference is required. Some important concepts are covered, but not in a complex way. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed.
- **Level 3 (Strategic thinking).** Deep knowledge becomes a greater focus at Level 3. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standards and items at Level 3 involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts.

2.2.6 Passage Types

Reading passages include both long and short texts selected from sources that students in each grade level would likely encounter in their classroom or in their independent reading. No passages were written specifically for the assessment but were instead collected from published works. Each passage is classified as described below.

- **Literary passages** are represented by a variety of genres—modern narratives; diary entries; drama; poetry; biographies; essays; excerpts from novels; short stories; and traditional narratives, such as fables, myths, and folktales.
- **Informational passages** are nonfiction and generally include two subgenres.
 - **Content passages** are primarily informational and often deal with the areas of science and social studies. They are drawn from sources such as newspapers, magazines, and books.
 - **Practical passages** are functional materials that instruct or advise the reader—for example, directions, reference tools, or reports.

The main difference among the passages used for grades 3 through 8 and 10 is their degree of complexity, which results from increasing levels of sophistication in language and concepts, as well as passage length. Measured Progress uses a variety of readability formulas to aid in the selection of passages

appropriate for the intended audience. In addition, the grade-level expertise of Montana teachers contributes to the selection of passages for each grade level.

Items related to these passages require students to demonstrate their skills in both literal comprehension, where the answer is stated explicitly in the text, and inferential comprehension, where the answer is implied by the text and/or the text must be connected to relevant prior knowledge to determine an answer. Items focus on the reading skills reflected in the content standards and require students to use reading skills and strategies to answer correctly—for example, how to identify the author’s principal purpose, such as to persuade, entertain, or inform—and to demonstrate their understanding of how words and images communicate to readers. Tables 2-4 and 2-5 depict passage distribution and length in grades 3 through 8 and 10.

**Table 2-4. 2010–11 Montana CRT:
Reading Passage Distribution Grades 3–8 and 10**

<i>Passage type</i>	<i>Passage content</i>	<i>Percent of test</i>	<i>Point distribution</i>
Literary	Stories, poetry, and other forms of literature	50%	30 points
Informational	Content and practical passages	50%	30 points
Total			60 points

<i>Passage length</i>	<i>Passage type</i>	<i>Percent of test</i>	<i>Point distribution</i>
Long	One literary or one informational per session	53%	32 points
Short	At least one literary and informational per session	47%	28 points
Total			60 points

**Table 2-5. 2010–11 Montana CRT:
Approximate Length of Reading Passages**

<i>Grade</i>	<i>Long passage (number of words)</i>	<i>Short passage (maximum word length)</i>
3	350–800	350
4	400–850	400
5	450–850	450
6	450–900	450
7	450–950	450
8	500–1,000	500
10	550–1,200	550

While every attempt is made to adhere to recommended grade-level word counts for long and short passages, the final decision to select a passage is based on extensive reviews by content area experts and bias panels, careful analysis of the sophistication of language and complexity of concepts in the passage, and the readability of the passage.

2.3 MATHEMATICS TEST SPECIFICATIONS

2.3.1 Standards

The mathematics specifications/blueprint is based on Montana's mathematics content standards:

- Mathematics Standard 1: Problem Solving
- Mathematics Standard 2: Numbers and Operations
- Mathematics Standard 3: Algebra
- Mathematics Standard 4: Geometry
- Mathematics Standard 5: Measurement
- Mathematics Standard 6: Data Analysis, Probability, and Statistics
- Mathematics Standard 7: Patterns, Relations, and Functions

2.3.2 Item Types

The Montana CRT in mathematics includes multiple-choice, short-answer, and constructed-response items. Short-answer items require students to perform a computation or solve a simple problem. Constructed-response items are more complex, requiring 8–10 minutes of response time. Each type of item is worth a specific number of points in the student's total mathematics score, as shown in Table 2-6.

Table 2-6. 2010–11 Montana CRT: Item Types

<i>Item type</i>	<i>Possible score points</i>
MC	0 or 1
SA	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; SA = short-answer; CR = constructed-response

2.3.3 Test Design

Table 2-7 summarizes the numbers and types of items that were used to construct the common portion of the Montana CRT in mathematics for 2010–11.

Table 2-7. 2010–11 Montana CRT: Common Mathematics Items

<i>Session</i>	<i>Calculator</i>	<i>Number of items grades 3, 4, 5</i>	<i>Calculator</i>	<i>Number of items grades 6, 7, 8, 9, 10</i>
1	Not Allowed	18 MC 2 SA 1 CR	Not Allowed	14 MC 3 SA 1 CR
2	Not Allowed	19 MC 1 SA	Allowed	21 MC
3	Allowed	18 MC 1 CR	Allowed	20 MC 1 CR

MC = multiple-choice; SA = short-answer; CR = constructed-response

2.3.4 Blueprints (Distribution of Points Across Standards)

Table 2-8 shows the distribution of points across the content standards.

**Table 2-8. 2010–11 Montana CRT:
Mathematics Specifications/Blueprint**

Content standards	<i>Raw score (percent point distribution by content strand)*</i>						
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Problem Solving and Number and Operations	34%(22)	34%(22)	32%(21)	32%(21)	27%(18)	27%(18)	20%(13)
Algebra	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	16%(11)
Geometry	15%(10)	15%(10)	16%(11)	16%(11)	18%(12)	18%(12)	20%(13)
Measurement	15%(10)	15%(10)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)
Data Analysis, Probability, and Statistics	12%(8)	12%(8)	15%(10)	15%(10)	18%(12)	18%(12)	20%(13)
Patterns, Relations, and Functions	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)

*Because percents are rounded to the nearest whole number, not all sums add up to 100%.

The mathematics test design consists of 55 multiple-choice items, three one-point short-answer items, and two four-point constructed-response items for 66 total points. There are two types of one-point items: multiple-choice and short-answer. The number of one-point items per strand will vary from year to year depending on which two strands are measured by the four-point items.

2.3.5 Depth of Knowledge

Each item on the Montana CRT in mathematics is assigned a depth of knowledge (DOK) level according to the cognitive demand of the item. Depth of knowledge is not synonymous with difficulty. The depth of knowledge level rates the complexity of the mental processing a student must use to solve a problem. Each of the levels is described below:

- **Level 1 (Recall).** This level requires the recall of a fact, definition, term, or simple procedure; the application of a formula; or the performance of a straight algorithmic procedure. Items at this level may require students to demonstrate a rote response.
- **Level 2 (Skill/Concept).** This level requires mental processing beyond that of a habitual response. These items often require students to make some decisions about how to approach a problem.
- **Level 3 (Strategic Thinking).** This level requires students to develop a plan or sequence of steps. These items are more complex and abstract than the items at the previous two levels. These items may also have more than one possible answer and may require students to use evidence, make conjectures, or justify their answers.

It is important that the Montana CRT in mathematics measure a range of depths of knowledge. Table 2-9 shows the percent and point ranges of the three depth of knowledge levels used on the CRT in mathematics.

**Table 2-9. 2010–11 Montana CRT:
Depth of Knowledge (DOK) Percent and Distribution by Level**

<i>DOK level</i>	<i>Percent range</i>	<i>Point range</i>
1	20% to 30%	13 to 20 points
2	60% to 75%	39 to 50 points
3	5% to 10%	4 to 8 points

2.3.6 Use of Calculators

Montana educators who helped develop the Montana CRT acknowledged the importance of mastering arithmetic algorithms. At the same time, they understood that the use of calculators is a necessary and important skill in society today. Calculators can save time and prevent error in the measurement of some higher-order thinking skills and allow students to do more sophisticated and intricate problems. For these reasons, calculators were permitted on some parts of the Montana CRT in mathematics and prohibited on other parts. (Students were allowed to use any calculator with which they were familiar.)

2.4 SCIENCE TEST SPECIFICATIONS

2.4.1 Standards

The science specifications are based on Montana’s science content standards:

- **Science Standard 1:** Scientific Investigations—Students, through the inquiry process, demonstrate the ability to design, conduct, evaluate, and communicate results and reasonable conclusions of scientific investigations.
- **Science Standard 2:** Physical Science—Students, through the inquiry process, demonstrate knowledge of properties, forms, changes, and interactions of physical and chemical systems.
- **Science Standard 3:** Life Science—Students, through the inquiry process, demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment.
- **Science Standard 4:** Earth/Space Science—Students, through the inquiry process, demonstrate knowledge of the composition, structures, processes, and interactions of Earth’s systems and other objects in space.
- **Science Standard 5:** Impact on Society—Students, through the inquiry process, understand how scientific knowledge and technological developments impact communities, cultures, and societies.
- **Science Standard 6:** Historical Development—Students understand historical developments in science and technology.

2.4.2 Item Types

The CRT in science includes multiple-choice and constructed-response items. Multiple-choice items require students to select the correct response from four choices, each item taking an average of one minute to answer. Constructed-response items are more involved, requiring 8–10 minutes of response time. Each type of item is worth a specific number of points in the student’s total science score, as shown in Table 2-10.

Table 2-10. 2010–11 Montana CRT: Item Types

<i>Item type</i>	<i>Possible score points</i>
MC	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; CR = constructed-response

2.4.3 Test Design

Table 2-11 summarizes the numbers and types of items that were used to compute student scores on the 2010–11 Montana CRT in science. Additionally, each test form had 13 multiple-choice matrix field-test items and one constructed-response field-test item that did not affect student scores.

Table 2-11. 2011–12 Montana CRT: Science Items

<i>Grades</i>	<i>Session 1</i>	<i>Session 2</i>	<i>Session 3</i>	<i>TOTAL</i>	
				MC	CR
4, 8, and 10	18 MC, 1 CR	17 MC	18 MC, 1 CR	53	2

MC = multiple-choice; SA = short-answer; CR = constructed-response

2.4.4 Blueprints (Distribution of Points Across Standards)

Table 2-12 shows the distribution of points and item types across the content standards.

Table 2-12. 2010–11 Montana CRT: Science Specifications/Blueprint

<i>Grades 4, 8, and 10</i>		
<i>Montana standards</i>	<i>Point distribution by content standard</i>	
	<i>Percent</i>	<i>Number</i>
1. Scientific Investigations	23%	14
2. Physical Science	23%	14
3. Life Science	23%	14
4. Earth/Space Science	23%	14
5. Impact on Society	8%	5
6. Historical Development	8%	5

The science test design consists of 53 multiple-choice items and two four-point constructed-response items for 61 total points. In any given year, the two constructed-response items will measure two different standards. From year to year, those standards may change.

2.4.5 Depth of Knowledge

Each item on the Montana CRT in science is assigned a depth of knowledge (DOK) level. The depth of knowledge level reflects the complexity of mental processing students use to answer an item. Depth of knowledge is not synonymous with difficulty. Each of the levels is described below.

- **Level 1 (Recall).** This level requires the recall of information such as a fact, definition, term, or simple procedure. These items require students only to demonstrate a rote response, use a well-known formula, or follow a set procedure.

- **Level 2 (Skill/Concept).** This level requires mental processing beyond that of recalling or reproducing a response. These items require students to make some decisions about how to approach the item.
- **Level 3 (Strategic Thinking).** This level requires reasoning, planning, and using evidence. These items require students to handle more complexity and abstraction than items at the previous two levels.

It is important that the Montana CRT in science measure a range of depths of knowledge. Table 2-13 shows the percent and point ranges of the three depth of knowledge levels used on the CRT in science.

**Table 2-13. 2010–11 Montana CRT:
Depth of Knowledge (DOK) Percent and Distribution by Level**

<i>DOK level</i>	<i>Percent range</i>	<i>Point range</i>
1	23% to 30%	14 to 18 points
2	51% to 67%	31 to 41 points
3	10% to 21%	6 to 13 points

2.4.6 Use of Calculators and Reference Sheets

Calculators are not used or needed when taking the science tests. There are no science reference sheets.

2.5 TEST DEVELOPMENT PROCESS

2.5.1 Item Development

Items used on the Montana CRT are developed and customized specifically for use on the Montana CRT and are consistent with Montana content standards, benchmarks, and grade-level expectations. Measured Progress test developers work with Montana educators to verify the alignment of items to the appropriate Montana content standards.

The development process combined the expertise of Measured Progress test developers and committees of Montana educators to help ensure items meet the needs of the CRT program. All items used on the common portions of the Montana CRT program were reviewed by a committee of Montana content area experts and by a committee of Montana bias experts. Tables 2-14 through 2-17 show the numbers of items developed within each content area for the 2010–11 Montana CRT.

Table 2-14. 2010–11 Montana CRT: Total Numbers of Items Developed by Content Area—Grades 3–8 and 10

<i>Grade</i>	<i>Mathematics</i>	<i>Reading</i>	<i>Science</i>
3	76	168	
4	76	168	78
5	76	168	
6	76	168	
7	76	168	
8	76	168	78
10	76	168	78

Table 2-15. 2010–11 Montana CRT: Annual Reading Item Development—Grades 3–8 and 10

<i>Passages</i>	<i>MC</i>	<i>CR</i>
2 long literary passages	40	4
2 long informational passages	40	4
4 short literary passages	40	0
4 short informational passages	40	0
12 total passages	160	8

MC = multiple-choice; CR = constructed-response

Table 2-16. 2010–11 Montana CRT: Annual Mathematics Item Development—Grades 3–8 and 10

<i>MC</i>	<i>SA</i>	<i>CR</i>
60	8	8

MC = multiple-choice; SA = short-answer; CR = constructed-response

Table 2-17. 2010–11 Montana CRT: Annual Science Item Development—Grades 4, 8, and 10

<i>MC</i>	<i>CR</i>
75	3

MC = multiple-choice; SA = short-answer; CR = constructed-response

Table 2-18 provides an overview of the item development process for common and field-test items, including the administration of the operational tests.

Table 2-18. 2010–11 Montana CRT: Item Development Process Overview

<i>Development step</i>	<i>Step details</i>
Selection of reading passages and external review for bias and sensitivity issues (December 2007)	Measured Progress test developers located potential reading passages. Reading passages were reviewed for bias and sensitivity issues before the development of reading items.
Development of items (January through May 2008)	Measured Progress test developers developed reading and mathematics items.
Item review for content appropriateness and for bias and sensitivity issues (May 2008)	Committees of Montana educators reviewed reading, mathematics, and science field-test items.
Editing of items (summer 2008)	Montana educators' recommended changes were incorporated into the items.
Field-test items (spring 2009)	Embedded field-test items were administered to a sample of students (minimum of 2,500 students per item) along with the 2009 operational test.
Item selection meeting (July 2009)	Measured Progress test developers and Montana educators selected common items for the spring 2010 operational CRT tests.
Operational test items (March 2010)	Items are part of the common item set and were used to determine student scores. Another embedded field test was also administrated.

2.5.2 Item Reviews at Measured Progress

A test developer within each content area reviewed items for

- item integrity, including content and structure, appropriateness to designated content area, format, clarity, possible ambiguity, and single correct answer.
- appropriateness and quality of reading selections and graphics.
- appropriateness of scoring guide descriptions and distinctions.
- whether the item is measuring the intended content standard.
- completeness of associated item documentation (e.g. scoring guide, content area codes, key, grade level, depth of knowledge, and contract identified).
- appropriateness for the designated grade level.

2.5.3 Item Reviews at State Level

All passages and items were reviewed in Montana. In December 2009, the Montana Passage Review Committee met to review passages that would be developed for the 2010–11 CRT administration. The committee consisted of teachers and education specialists from across the state. In April 2010, Montana educators from across the state reviewed field-test items for content appropriateness, alignment to standards, depth of knowledge, and grade-level appropriateness.

2.5.4 Bias and Sensitivity Review

Bias review is an essential component of the development process. During the bias review process, reading passages and items from all content areas were reviewed by a committee of Montana educators. Items were examined for issues that might offend or dismay students, teachers, or parents. Including such groups in

the development of assessment items and materials can avoid many controversial issues, and concerns can be allayed before the test forms are produced.

2.5.5 Reviewing and Refining

Recommended changes from the Item Review and Bias and Sensitivity meetings were incorporated into the items by Measured Progress test developers.

2.5.6 Item Editing

Measured Progress editors then reviewed and edited the items to ensure adherence to sound testing principles and to style guidelines in the *Chicago Manual of Style, 15th ed.* These principles include the stipulations that items

- demonstrate correct grammar, punctuation, usage, and spelling;
- are written in a clear, concise style;
- contain unambiguous explanations that tell students what is required to attain a maximum score;
- are written at a reading level that allows students to demonstrate their knowledge of the subject matter being tested regardless of reading ability;
- exhibit high technical quality regarding psychometric characteristics;
- have appropriate answer options or score-point descriptors; and
- are free of potentially insensitive content.

2.5.7 Item Selection and Operational Test Assembly

In July 2010, Measured Progress test developers met with Montana educators to select common items. In preparation for the meeting, the test developers and psychometricians at Measured Progress considered the following in selecting sets of items to propose for the common item set to be used on the 2011 assessment:

- **Content coverage/match to test design and blueprints.** The test designs and blueprints stipulate a specific number of multiple-choice and constructed-response items for each content area. Item selection for the embedded field test was based on the number of items in the existing pool of items that are eligible for the common.
- **Item difficulty and complexity.** Item statistics drawn from the data analysis of previously field-tested items were used to ensure similar levels of difficulty and complexity from year to year as well as quality psychometric characteristics.
- **“Cueing” items.** Items were reviewed for any information that might “cue” or provide information that would help to answer another item.

At the meeting, the Montana educators reviewed the proposed sets of items and made the final selection of items for the common.

The test developers then sorted and laid out the items into test forms. During assembly of the test forms, the following criteria were considered:

- **Key patterns.** The sequence of keys (correct answers) was reviewed to ensure that their order appeared random.
- **Option balance.** Items were balanced across forms so that each form contained a roughly equivalent number of key options (As, Bs, Cs, and Ds).
- **Page fit.** Item placement was modified to ensure the best fit and arrangement of items on any given page.
- **Facing-page issues.** For multiple items associated with a single stimulus (reading passages) and multiple-choice items with large graphics, consideration was given to whether those items needed to begin on a left- or right-hand page and to the nature and amount of material that needed to be placed on facing pages. These considerations served to minimize the amount of page flipping required of students.
- **Relationships among forms.** Although field-test items differ from form to form, these items must take up the same number of pages in all forms so that sessions begin on the same page in every form. Therefore, the number of pages needed for the longest form often determines the layout of each form.
- **Visual appeal.** The visual accessibility of each page of the form was always taken into consideration, including such aspects as the amount of “white space,” the density of the test, and the number of graphics.

2.5.8 Operational Test Draft Review

After the forms were laid out as they would appear in the final test booklets, the forms were again thoroughly reviewed by Measured Progress editors to ensure that the items appeared exactly as intended. Any changes made during test construction were reviewed and approved by the test developer.

2.5.9 Alternative Presentations

Form 1 for the grades 3 through 8 and 10 tests was translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind and visually impaired students. In addition, Form 1 for each grade was adapted into a large-print version.

2.6 TEST SESSIONS

The Montana CRT was administered during the spring of 2011 during a four-week period from March 1, 2011 to March 24, 2011. Reading and mathematics tests were administered in grades 3 through 8 and 10, and science tests were administered in grades 4, 8, and 10. Schools were able to schedule testing sessions at any time during the four-week period, provided they followed the sequence detailed in the scheduling guidelines in the *Test Administrator's Manual*. Schools were asked to schedule makeup tests for students who were absent from initial test sessions during the testing window.

CHAPTER 3. TEST ADMINISTRATION

3.1 RESPONSIBILITY FOR ADMINISTRATION

As indicated in the *Test Coordinator's Manual*, principals and/or their designated school test coordinators are responsible for the proper administration of the CRT. This report was used to ensure the uniformity of administration procedures from school to school.

3.2 ADMINISTRATION PROCEDURES

School test coordinators were instructed to read the *Test Coordinator's Manual* prior to testing and to be familiar with the instructions given in the *Test Administrator's Manual*. The *Test Coordinator's Manual* provides each school with checklists to help prepare for testing. The checklists outline tasks to be performed before, during, and after test administration. In addition to providing these checklists, the *Test Coordinator's Manual* outlines the nature of the testing materials sent to each school, how to inventory the materials, how to track the materials during administration, and how to return the materials once testing was complete. The *Test Coordinator's Manual* also contains information about including or excluding students. The *Test Administrator's Manual* includes checklists for administrators to prepare themselves, their classrooms, and their students for administration of the test. The *Test Administrator's Manual* contains sections that detail the procedure to be followed for each test session, and it contains instructions for preparing the materials prior to giving them to school test coordinators for return to Measured Progress.

The Montana CRT is an untimed assessment; however, guidelines or ranges were provided in the *2011 Test Coordinator's Manual* and the *2011 Test Administrator's Manual* based on the following estimates of the time it takes an average student to respond to each type of item on the test:

- Multiple-choice items—1 minute per item
- Short-answer items—2 minutes per item
- Constructed-response items—10 minutes per item

The provided guidelines suggested scheduling 45–55 minutes per test session (50–60 minutes for grade 10 students). The guidelines also suggested scheduling a break between each of the three sessions in each content area to prevent fatigue.

While the guidelines for scheduling were based on the assumption that most students would complete the test within the estimated amounts of time, each test administrator was asked to allow additional time for students who needed additional time to complete the test. If additional classroom space was not available for this purpose, schools were encouraged to use another space, such as a guidance office. If other areas were not available, the guidelines recommended scheduling each classroom used for test administration for the maximum possible amount of time.

3.3 PARTICIPATION REQUIREMENTS AND DOCUMENTATION

All students were expected to participate in the CRT; however, the scores of students in the following categories were excluded from the calculation of averages:

- foreign exchange students
- students not enrolled in an accredited Montana school (for example, home-schooled students)
- students enrolled in a private accredited school
- students enrolled in a private non-accredited school
- students enrolled in a private non-accredited Title 1 school
- students enrolled part-time (less than 180 hours) taking a mathematics or reading course
- first year in U.S. LEP students who were required to participate in the mathematics assessment only
- students who took the CRT using a “nonstandard” accommodation

A summary of this information is shown in Table 3-1, which was published in the *Test Administrator’s Manual* and the *Test Coordinator’s Manual*.

Table 3-1. 2010–11 Montana CRT: Summary of Eligibility for Exclusion from the CRT

<i>Excluded from averages</i>	<i>MUST participate</i>	<i>MAY participate</i>
Foreign exchange students	Yes	
Students not enrolled in an accredited Montana school		Yes
Students enrolled in a private accredited school	Yes	
Students enrolled in a private non-accredited school		Yes
Students enrolled in a private non-accredited Title I school		Yes
Students enrolled part-time (less than 180 hours) taking a mathematics or reading course		Yes
Reading: first year in U.S. LEP students		Yes
Mathematics: first year in U.S. LEP students	Yes	

Staff members coded information about exclusion, if applicable; in the student response booklet after testing was completed. The *Test Coordinator’s Manual* and *Test Administrator’s Manual* provide detailed instructions for coding exclusions and accommodations. In addition, testing exclusions were discussed thoroughly in the pre-administration training audio CD (see Appendix A: Reporting Decision Rules).

A summary of participation on the 2010–11 Montana CRT by demographic category for each content area is shown in Appendix B.

3.3.1 Students with Disabilities

All students with special needs participate in the CRT assessment program, either by taking the regular CRT or CRT-Alternate Assessment if they meet the eligibility criteria.

Form 1 for the grades 3 through 8 and 10 tests was enlarged to 18-point font for visually impaired students and was translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind students. Students with special needs and LEP students are often given these test accommodations.

3.4 ADMINISTRATOR TRAINING

The OPI hosted a test-administration workshop in Billings, Montana, on January 20–21, 2011. The workshop was well attended, but attendance of system and school test coordinators was not mandatory. OPI and Measured Progress staff members hosted six sessions that covered test accommodations, student information system (AIM) updates, CRT materials and administration, CRT-Alternate materials and administration, online reporting, and test security.

In addition to the workshop and the distribution of the *2011 Test Coordinator's Manual* and *Test Administrator's Manual*, the OPI and Measured Progress produced the PowerPoint presentation, "Spring 2011: CRT and CRT-ALT Overview and Update of System and School Test Coordinators." Training materials and the audio PowerPoint presentation were posted on the OPI's Web site: <http://www.opi.mt.gov>. The training power point presentations provided the training information for system and school test coordinators who were unable to attend the administration workshops. The power point presentations also served as useful tools for training both system and school personnel.

3.5 DOCUMENTATION OF ACCOMMODATIONS

The 2011 CRT Accommodations Manual and the accommodations training PowerPoint, *Guidelines and Procedures for CRT Accommodations*, were produced by the OPI and were included on the CRT training CD provided to each system and school in the first shipment received by systems in early February 2011 from Measured Progress. General instructions regarding accommodation usage and a list of available accommodations were included in the *2011 Test Coordinators Manual*.

Standard accommodations were available to all students on the basis of individual needs and regardless of disability status. Decisions regarding standard accommodations were made by the student's educational team on an individual basis, consistent with either previous accommodation decisions for the student or current educational needs. Accommodations usage was to be consistent with those used during the student's regular classroom instruction and assessment for at least three months prior to testing.

Nonstandard accommodations were offered to students with disabilities only if the accommodations are specified in the student's IEP. If a student was assessed with a nonstandard accommodation, the student was considered a nonparticipant when calculating the participation rate for AYP purposes. In addition to the student being considered a nonparticipant, the student's score from the assessment is not included in calculating the proficiency rate for AYP.

Table 3-2 below shows the number of students at each subject and grade who were tested with and without accommodations. In addition, frequencies of accommodations used by accommodation type are presented in Appendix C.

Table 3-2. Number of Students Tested With and Without Accommodations by Subject and Grade

Subject	Grade	Number of students tested:	
		With accommodations	Without accommodations
Mathematics	3	1,767	8,725
	4	1,849	8,696
	5	1,578	8,790
	6	1,383	9,139
	7	1,157	9,325
	8	1,112	9,336
	10	632	9,427
Reading	3	1,698	8,764
	4	1,754	8,770
	5	1,554	8,802
	6	1,354	9,169
	7	1,130	9,365
	8	1,131	9,353
	10	655	9,422
Science	4	1,664	8,911
	8	1,091	9,392
	10	669	9,408

3.6 TEST SECURITY AND ADMINISTRATION IRREGULARITIES

Test coordinators and administrators are prohibited from disclosing the contents of CRT assessments.

Under no circumstances should test booklets or marked answer booklets be circulated among faculty, administrators, or other persons.

All system test coordinators and school principals received the *OPI Guidelines and Procedures for Test Security*. This OPI publication was made available to system superintendents, principals, and test administrators for the purpose of outlining the reporting procedures for security and administration violations. All concerns about breaches of test security or noncompliance with test administration procedures were to be reported immediately to the principal, system test coordinator, and state assessment director.

3.7 TEST ADMINISTRATION WINDOW

The Montana CRT was administered during the spring of 2011 during a four-week period from March 1, 2011 to March 24, 2011. Reading and mathematics tests were administered in grades 3 through 8 and 10, and science tests were administered in grades 4, 8, and 10. Schools were able to schedule testing sessions at any time during the four-week period, provided they followed the sequence detailed in the scheduling guidelines in the *Test Administrator's Manual*. Schools were asked to schedule makeup tests for students who were absent from initial test sessions during the testing window.

3.8 SERVICE CENTER

To address testing concerns, Measured Progress established a help desk dedicated to the Montana CRT. Service Center support is an essential element to the successful administration of large-scale assessments. It provides a central location that individuals in the field can call via a toll-free number to request assistance, report problems, or ask specific questions.

The Measured Progress help desk provided support during all phases of the testing window. It was staffed at varying levels, based on need and the volume of calls received, from 8:00 a.m. to 4:00 p.m. MST. At a minimum, the help desk consisted of a product support specialist responsible for receiving, responding to, and tracking calls and e-mails, and routing issues to the appropriate person(s) for resolution. In addition, the program manager and/or program assistant addressed communications that required a higher level of program support.

During the period between February 22, 2011, when the testing materials were delivered to schools, and April 6, 2011 when the materials were returned to Measured Progress, the Service Center received 102 calls. The majority of these calls were to order additional materials for students who enrolled after materials were shipped and to arrange for UPS to pick up the materials after testing. The service center staff also responded to administration questions and referred policy questions regarding test security or accommodations usage to the OPI.

CHAPTER 4. SCORING

Accurate and timely scoring of constructed-response (CR), short-answer (SA), and multiple-choice (MC) items is an important process in any successful assessment program. This chapter defines the scope and processes of Measured Progress's scoring services for the 2010–11 Montana CRT.

4.1 MACHINE-SCORED ITEMS

Preceding the arrival of the Montana CRT student response booklets, Measured Progress prepared customized scanning programs to enable selective reading of all scannable materials including student identification and demographics and to electronically format the scanned information.

Once the student answer documents were received from each Montana school following test administration, Measured Progress optically scanned each page from every student booklet to create digital images of the entire document. Every page was bar coded so that the scores applied to each item could be linked to the correct student, school, and district. Student responses were then imported into iScore™, Measured Progress's proprietary, image-based scoring system, for secure processing and scoring. By using this image scoring system, Measured Progress was able to increase reliability and productivity as well as monitor and maintain quality control.

Student multiple-choice response data was machine-scored at the same time that student constructed-response and short-answer items were scanned into iScore for person-scoring. Multiple-choice items were compared to scoring keys via item analysis software. Correct multiple-choice answers were assigned a score of one point and incorrect answers were assigned zero points. Student multiple-choice responses consisting of multiple marks and blank responses were also assigned zero points.

Student responses that could not physically be scanned (e.g., documents damaged during administration or shipment) were physically reviewed and scored on an individual basis by trained, qualified staff. These scores were linked to the student's demographic data and merged with the student's scoring file by Measured Progress's data processing department.

Table 4-1. 2010–11 Montana CRT Number of Responses Scanned and Scored

<i>Content area</i>	<i>Grade</i>	<i>Number of responses scanned and scored</i>
Mathematics	3	54,431
	4	55,036
	5	53,911
	6	54,660
	7	54,400
	8	54,670
	10	53,186
Reading	3	22,100
	4	22,260
	5	21,810
	6	22,110
	7	22,114
	8	21,116
	10	21,520
Science	4	22,260
	8	22,116
	10	21,520

4.2 PERSON-SCORED ITEMS

Scanned images of open-response items were processed and organized into item-specific groups in preparation for person-scoring by iScore. iScore’s secure, Web-based application provided qualified staff, including readers and their leadership, and password-protected access for reading and scoring electronic student responses at one or multiple scoring sites without compromising confidentiality. The digital image clip information of constructed-response and short-answer responses allowed iScore to replicate student responses just as they appeared on the originals and to display the replicated responses on individual monitors for person-scoring. In addition, the processes of item benchmarking, reader training, scoring, editing/cleanup, and reporting were all accomplished electronically and without further reference to the originals.

Organized by iScore in this way, qualified readers were able to view only one response from a single item at a time. Because item responses were tracked and distributed among groups of readers by iScore, each response in an individual student’s response booklet could be assigned to and scored by a different reader. This maximization of the number of readers per student response booklet effectively minimized bias errors caused by reader sampling.

Leadership staff, on the other hand, had constant, albeit view-only, access to all of the imaged responses from a student’s booklet for whenever necessary. The actual test booklets and answer documents were also available to the content area chief reader and the iScore operational manager (see section on “Scoring Location and Staff”).

To ensure the security of constructed-response and short-answer items and responses scored, all scoring activities in iScore were performed “blind,” i.e., without student names, district, and/or school

information visible or able to be associated with responses or raw scores. During scoring, iScore distributed images of student responses to the computer monitors of readers located at one of Measured Progress's scoring facilities. When iScore sent an image of a student response to an individual reader's computer monitor, the reader evaluated the response and recorded the score via keypad or mouse entry. Once the score was entered, a new response appeared immediately on the screen.

Although iScore is based on conventional, best practice scoring procedures, it also offers the following benefits:

- It provides leadership staff with real-time information about group and individual level performance including scoring accuracy and consistency as well as overall process monitoring and reporting.
- It ensures the randomized distribution of student responses among readers during scoring and automatically assigned student responses to one or more scorers for interrater agreement monitoring.
- It permits password-only access limited to those solely in the employ of Measured Progress and working within a qualified scoring or scoring management capacity.
- It maintains student anonymity and confidentiality by masking student biographical information from viewers.
- It offers immediate access to samples of student responses and scores for reporting and analysis.
- It offers early access to subsets of data for tasks such as standard setting.
- It reduces material handling, which saved time and labor while enhancing the security of materials.

The iScore database, its control operation, and its administrative offices were all based in Dover, New Hampshire. The iScore system monitored accuracy, reliability, and consistency across all Measured Progress scoring facilities. To ensure that scoring information and updates were equally shared and implemented across all scoring facilities, constant communication and coordination was accomplished daily via e-mail, telephone, fax, and secure, Web-based networks.

4.2.1 Scoring Location and Staff

Scoring Location

Scoring the 2010–11 Montana CRT program took place at Measured Progress's scoring facilities located in Louisville, Kentucky, and Menands, New York. The overview of scoring operations is presented by content area and grade in Table 4-2.

Table 4-2. 2010–11 Montana CRT Scoring Locations by Content and Grade

<i>Content / grade level</i>	Louisville, KY (1 st shift)	Menands, NY (2 nd shift)
Mathematics Grade 3	X	
Mathematics Grade 4	X	
Mathematics Grade 5	X	
Mathematics Grade 6	X	
Mathematics Grade 7	X	
Mathematics Grade 8	X	
Mathematics Grade 10	X	
Reading Grade 3	X	
Reading Grade 4	X	
Reading Grade 5	X	
Reading Grade 6	X	
Reading Grade 7	X	
Reading Grade 8	X	
Reading Grade 10	X	
Science Grade 5		X
Science Grade 8		X
Science Grade 10		X

Scoring Staff

Staffing for the 2010–11 Montana CRT implemented low scoring-leadership-to-reader ratios and was composed of the following Measured Progress staff members:

- Scoring project manager, who oversaw the overall contract from a scoring perspective and acted as a liaison with contract management staff, data analysis staff, and the client while managing the content area experts (chief readers, quality assurance coordinators, etc.).
- Chief readers, who prepared benchmarking/training materials and led the review and client approval of materials, working closely together with Measured Progress development specialists and Montana educators. Chief readers trained, qualified, and monitored readers during the scoring process; supervised quality assurance coordinators, senior readers, and readers; and monitored scoring accuracy and consistency. The ratio of chief readers to the scoring project manager was 3:1.
- Quality assurance coordinators (QACs), who managed the training and benchmarking of grades and items within the Montana CRT. QACs trained, qualified, and monitored readers during the scoring process; supervised senior readers and readers; and monitored scoring accuracy and consistency. The ratio of QACs to chief readers was 7:1.
- Senior readers (SRs), who supervised readers during the scoring process and monitored scoring accuracy and consistency while managing quality control measures via iScore. The ratio of SRs to QACs was 1:1.

- Readers, who were qualified, temporary staff members performing the bulk of scoring work, evaluating and scoring student responses according to the Montana CRT guidelines provided for each grade level and content area scored. Readers received the same orientation and training as direct hires. The ratio of readers to SRs was 9:1.

4.2.2 Reader Recruitment and Qualifications

In preparation for scoring the 2010–11 Montana CRT, Measured Progress actively sought and recruited readers to represent a diverse spectrum of educational, professional, and ethnic populations. The customary cross-section of readers employed included content area specialists such as editors, business professionals, scientists, authors, graduate school students, and both current and retired educators.

Although the employment of readers holding a four-year college degree or higher was preferred, all readers were required to have successfully completed a minimum of at least two years of college and to have demonstrated knowledge of the content area they scored. All readers were required to submit documentation (i.e., college transcript and/or resume) of their qualifications.

For training and qualification, readers were placed at grade levels and in content areas that matched their areas of experience and expertise. Reader demographic information (gender, educational, and ethnic background, etc.) was electronically documented for reporting. All readers were subject to stringent nondisclosure requirements and supervision and were required to sign a nondisclosure/confidentiality agreement. Table 4-3 summarizes the educational credentials of the 2010–11 Montana CRT readers and QACs.

Table 4-3. 2010–11 Montana CRT Education Credentials of Readers and QACs

Readers				
Description	Louisville, KY (1 st shift)	Menands, NY (2 nd shift)	Total	Percent
Less than 48 college credits	0	0	0	0.00%
48+ college credits	14	1	15	6.47%
Associate's degree	17	9	26	11.21%
Bachelor's degree	88	19	107	46.12%
Master's degree	57	17	74	31.90%
Doctorate	8	2	10	4.31%
Total	184	48	232	100.00%
Scoring leadership				
Description	Louisville, KY (1 st shift)	Dover, NH (2 nd shift)	Total	Percent
Less than 48 college credits	0	0	0	0.00%
48+ college credits	3	0	3	8.57%
Associate's degree	1	0	1	2.86%
Bachelor's degree	15	4	19	54.29%
Master's degree	5	3	8	22.86%
Doctorate	4	0	4	11.43%
Total	28	7	35	100.00%

4.2.3 Methodology for Scoring Polytomous Items

Possible Score Points

The ranges of possible score points for the different polytomous items found on the 2010–11 Montana CRT are shown in Tables 4-4 and 4-5.

Table 4-4. 2010–11 Montana CRT Short-Answer Item Scoring Guide

<i>Score point</i>	<i>Description</i>
1	The student's response provides a complete and correct answer.
0	The student's response is totally incorrect or too minimal to evaluate.
B	Blank/no response.

Table 4-5. 2010–11 Montana CRT Constructed-Response Item Scoring Guide

<i>Score point</i>	<i>Description</i>
4	The student completes all important components of the task and communicates ideas clearly.
	The student demonstrates in-depth understanding of the relevant concepts and/or processes.
	When instructed to do so, the student chooses more efficient and/or sophisticated processes.
	When instructed to do so, the student offers insightful interpretations or extensions (e.g., generalizations, applications, and analogies).
3	The student completes the most important components of the task and communicates clearly.
	The student demonstrates understanding of major concepts even though he/she overlooks or misunderstands some less important ideas or details.
2	The student completes most important components of the task and communicates those clearly.
	The student demonstrates that there are gaps in his/her conceptual understanding.
1	The student shows minimal understanding.
	The student addresses only a small portion of the required task(s).
0	The student's response is totally incorrect or irrelevant.
B	Blank/no response.

Condition Codes

When numerical score-point parameters did not apply to a student response, readers had the option of designating one the following options:

- Blank response (empty entry without an attempt at responding to the question)
- Unreadable response (response is illegible or too faint to accurately interpret)
- Wrong Location (a relevant response entered into the space reserved for a different item)
- Non-English response (a response written entirely in a language other than English)

Unreadable and Wrong Location responses were resolved by consulting the original test booklet and/or by identifying the correct location.

4.2.4 Reader Training

For each item scored in the 2010–11 Montana CRT, Measured Progress readers were required to demonstrate their scoring ability by participating in training sessions specific to each student response item scheduled to be scored. The scoring project began with an introduction of the onsite scoring staff and an overview of the Montana CRT program’s purpose and goals (including discussion about document security, student confidentiality, the proprietary nature of testing materials, scoring materials, and iScore procedures).

Actual training began with groups of readers organized into content area-, grade-, and item-specific group assignments. Each reader was provided a personal hard copy of item-specific training materials distributed at the beginning of each work session and had to account for these materials during secure collection at the end of each work session. During training, readers were strongly encouraged to take notes and highlight their own hard copies of the training materials.

For each item trained, the QAC assigned to the item commenced reader training by reviewing and discussing the item-specific scoring guide. The training QAC demonstrated the process of applying the item’s scoring guide and score point descriptors to the exemplars found in the subsequent *anchor* and *training sets* before attempting to demonstrate scoring accuracy in the *qualifying set*.

Anchor Set

This is a set of responses approved by the respective content area specialists for reading, mathematics, or science representing the OPI. Each anchor set contained one OPI-approved sample response per score point considered to be a mid-range exemplar. This set occasionally included a second sample response if there was more than one plausible way to illustrate the merits and intent of a score point. Responses in the anchor sets were typical, rather than unusual or uncommon; solid, rather than controversial or borderline; and true,

meaning that their scores could not be changed except by the OPI and Measured Progress test development specialists.

Each anchor set response was read aloud to readers by the training QAC. Training QACs facilitated group discussion of responses in relation to the scoring guide and score point descriptors to help readers internalize the typical characteristics of score points. The anchor set served as a reference for readers as they went on to scoring sample responses in the training set that followed.

Training Set

Next, readers practiced applying the scoring guide and anchor set to responses in the training set. The training set typically included 6 to 10 student responses designed to help establish both the full score point range and the variation of possible responses within each score point. The training set often included unusual responses that were less clear or solid (e.g., briefer than normal, employing atypical approaches, simultaneously containing very low and very high attributes, and written in ways difficult to decipher).

Responses in the training set were presented to readers without scores and in a randomized score point order. Once readers had independently read and applied their score to a training set response, the training QAC would poll readers and then announce the actual response score. The QAC then responded to reader questions and/or comments while pointing out particular scoring issues at hand (e.g., the borderline between two score points). Throughout each item training, the QAC continually routed reader discussion of score points back to the anchor set and scoring guide. After the training set had been completed, readers were required to demonstrate scoring accuracy using *qualifying sets* assembled for constructed-response items.

Qualifying Set

Following participation in each item training session, readers were administered a qualification set of committee-reviewed responses (CRRs) assigned to each item in the reader's content area. Each qualifying set was composed of ten preselected, previously scored responses chosen as clear illustrations of score point examples that would measure which readers had adequately internalized item training before those readers were able to score live student responses. These CRRs were selected by scoring leadership and randomly distributed to each reader via iScore during qualification.

In order to qualify on a qualification set, readers were required to demonstrate a scoring accuracy level of at least eighty percent (80%) exact agreement (i.e., exactly match scores on at least 8 of the 10 CRRs) and at least ninety percent (90%) exact-or-adjacent agreement (i.e., exactly match or be within one score point on 9 or 10 of the 10 CRRs). In other words, readers were allowed 1 discrepant score (i.e., 1 score out of the 10 CRRs that was more than one score point from the CRR score) provided they had at least 8 exact scores.

Once a group of readers successfully qualified on a particular item, responses for that item in iScore were randomly assigned and presented to them on their computer monitors, one response at a time. Readers

unable to qualify on the first qualification set received QAC retraining (see section on “Retraining”) and a subsequent opportunity at qualification on a second qualification set. Readers unable to qualify on the second qualification set were not eligible to score that item.

(Note: In order to be eligible to score short-answer mathematics items in grades 3 through 8 and 10, readers were required to qualify on at least one mathematics constructed-response item for that grade.)

Retraining

Readers unable to qualify on the first qualification set received QAC retraining by reviewing their performance in relation to the item training materials. The QAC responded to reader questions and routed discussion of score points back to the anchor set and scoring guide. Readers were then allowed the opportunity at qualification on a second qualification set. Readers unable to qualify on the second qualification set were not eligible to score that item. Table 4-6 depicts the accuracy and qualification percentages of the reader applicants.

Table 4-6. 2010–11 Montana CRT Scoring Accuracy and Qualification Statistics

<i>Content area</i>	<i>Grade</i>	<i>Item</i>	<i>Average percent exact agreement</i>		<i>Readers</i>		
			<i>Embedded CR sets</i>	<i>Double-blind scoring</i>	<i>Taking qualification sets</i>	<i>Successfully qualifying</i>	<i>Percent qualifying</i>
Mathematics	3	23	NA	97.4	NA	NA	
		24	NA	98.3	NA	NA	NA
		25	92.1	85.4	16	15	93.8
		48	NA	96.6	NA	NA	NA
		72	90.0	85.3	13	12	92.3
	4	23	NA	97.0	NA	NA	NA
		24	NA	98.7	NA	NA	NA
		25	88.8	90.4	13	13	100.0
		48	NA	97.9	NA	NA	NA
		72	91.2	85.1	13	13	100.0
	5	23	NA	97.5	NA	NA	NA
		24	NA	97.3	NA	NA	NA
		25	79.3	79.2	12	12	100.0
		48	NA	97.0	NA	NA	NA
		72	95.0	89.9	13	12	92.3
	6	18	NA	97.8	NA	NA	NA
		19	NA	98.0	NA	NA	NA
		20	NA	99.3	NA	NA	NA
		23	93.9	95.7	13	13	100.0
		73	80.7	84.8	12	12	100.0
	7	18	NA	98.3	NA	NA	NA
		19	NA	98.2	NA	NA	NA
		20	NA	97.9	NA	NA	NA
		23	88.9	93.9	11	11	100.0
		73	88.3	88.0	12	12	100.0
	8	18	NA	98.1	NA	NA	NA

continued

Content area	Grade	Item	Average percent exact agreement		Readers		
			Embedded CR sets	Double-blind scoring	Taking qualification sets	Successfully qualifying	Percent qualifying
Reading		19	NA	95.4	NA	NA	NA
		20	NA	96.9	NA	NA	NA
		23	85.0	91.6	19	19	100.0
		73	89.0	89.1	17	16	94.1
	10	18	NA	96.1	NA	NA	NA
		19	NA	95.4	NA	NA	NA
		20	NA	96.4	NA	NA	NA
		23	80.9	93.5	14	14	100.0
		73	90.3	93.7	14	14	100.0
	3	27	75.7	80.8	13	12	92.3
		81	73.3	79.4	17	14	82.4
	4	27	87.5	80.8	13	13	100.0
		81	67.1	75.9	13	12	92.3
	5	27	70.1	68.7	13	12	92.3
		81	74.6	73.3	11	9	81.8
	6	27	79.7	79.6	13	13	100.0
		81	78.5	72.0	11	11	100.0
	7	27	70.3	69.7	13	12	92.3
		81	76.7	73.3	16	15	93.8
	8	27	90.2	66.4	13	13	100.0
		81	88.1	73.2	20	19	95.0
	10	27	82.4	79.0	13	13	100.0
		81	81.7	84.4	11	11	100.0
Science	4	23	92.6	77.5	15	15	100.0
		69	86.0	81.7	13	12	92.3
	8	23	89.0	86.6	16	16	100.0
		69	97.7	94.2	16	16	100.0
	10	23	89.2	82.4	20	20	100.0
		69	84.9	89.0	20	20	100.0

4.2.5 Leadership Training

A core group of scoring leadership staff including QACs and SRs guided and monitored readers throughout the scoring process. Because quality control by QACs and SRs moderated the scoring process and maintained the integrity of scores, the individuals chosen to fill those positions were selected for their accuracy and consistency. In addition, the training QACs assigned to train readers were also selected for their ability to instruct and content area specialization.

The purpose of leadership training was to ensure that QACs provided thorough and consistent training and feedback to readers. Chief readers were able to discuss item details and score point rationale within training materials in order to prepare scoring leadership for reader questions. Chief readers trained and reviewed items with QACs, who in turn trained and reviewed items with their SRs. During actual item

scoring, QACs trained and supervised readers and monitored SR accuracy and consistency. The SRs, in turn, supervised their own group of readers and monitored reader accuracy and consistency. Similar to readers, scoring leadership who performed quality control measures in iScore were also required to qualify using qualification sets by demonstrating a scoring accuracy level of at least eighty percent (80%) exact agreement and at least ninety percent (90%) exact-or-adjacent agreement.

4.2.6 Monitoring of Scoring Quality Control

iScore was preprogrammed to monitor individual reader accuracy and scoring consistency among readers on a constant basis. iScore's use of multiple monitoring techniques was critical to the process of live scoring, allowing readers who met or exceeded accuracy standards to continue scoring. Reader accuracy and consistency was measured in iScore throughout the scoring process using the following methods and tools:

- Embedded Committee-Reviewed Responses (CRRs)
- Read-Behind Scoring
- Double-Blind Scoring
- Reader Arbitration

Embedded Committee-Reviewed Responses

Embedded Committee-Reviewed Responses (Embedded CRRs) are preselected, previously scored responses used to ensure that readers had adequately internalized item training and remained calibrated to the scoring standard during live scoring. Previous to scoring, scoring leadership selected Embedded CRRs for each item and loaded the examples into iScore ("embedded"). Each example represented images of actual student work and appeared no different from live student responses. During the first day of live scoring of each item, iScore randomly distributed thirty (30) Embedded CRRs to each reader. Embedded CRRs were employed for all constructed-response items and enabled scoring leadership to monitor reader accuracy and consistency as gauged by the known scores of the Embedded CRRs.

Readers with a disproportionate number of adjacent and/or discrepant scores in Embedded CRRs were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds.

Read-Behind Procedures

Read-behind scoring refers to scoring leadership (typically a SR) scoring a response that was recently scored by a reader. The gain was an immediate, real-time snapshot of each reader's accuracy and consistency during scoring. SRs were required to perform read-behinds on each reader throughout each day and at any point during scoring. This practice was applied to all open-ended item types and performed on all readers as required.

Once called up in iScore by the SR, read-behind responses were selected by iScore and placed into the SR's read-behind queue. Readers were aware neither of iScore's selection nor which of their scored responses was to be reviewed by their SR. Likewise, SRs were not aware of the reader's score when iScore presented each read-behind response for their own review and eventual score. The SR then applied their own score to the response before the reader's score was made viewable in iScore. This SR review and comparison of the two scores created the score of record determination (i.e., the reported score) as follows:

- If the reader and SR applied the same score (exact agreement), no action was necessary; the reader's score became the score of record.
- If the reader and SR scores differed by 1 point (adjacent), the SR's score became the score of record, thereby overriding the reader's score.
- If the reader and SR scores differed by more than 1 point (discrepant), the SR's score became the score of record, thereby overriding the reader's score.

Readers with a disproportionate number of adjacent and/or discrepant scores in read-behinds were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds. Table 4-7 outlines the resolution of reader scores using the read-behind procedure.

Table 4-7. 2010–11 Montana CRT Examples of Read-Behind Scoring Resolution

<i>Reader score</i>	<i>QAC/SR score</i>	<i>Score of record</i>
4	4	4
4	3	3*
4	2	2*

* QAC/SR's score.

Double-Blind Scoring

Scoring procedures for both constructed-response and short-answer item types included both single-scoring and double-scoring. Single-scored responses were scored by one reader. Double-scored responses were scored “blindly” by two different readers, unaware of the other’s score. These double-blind scores were monitored for interrater agreement accuracy and scoring consistency. A default minimum setting of two percent (2%) from all open-ended item types was double-blind scored. In addition, responses marked blank, unreadable, and non-English were automatically routed for double-blind scoring. Table 4-8 indicates the frequency for which open-ended item responses from each content area were double-blind scored.

Table 4-8. 2010–11 Montana CRT Frequency of Double-Blind Scoring by Grade and Content

<i>Grade</i>	<i>Content area</i>	<i>Responses double-blind scored</i>
3–8, 10	Mathematics	2% minimum
3–8, 10	Reading	2% minimum
4, 8, 10	Science	2% minimum
All	Unreadable responses	100%
All	Blank responses	100%
All	Non-English	100%

Reader Arbitration

When double-blind scores applied by two readers on a single response differed by more than 1 point (a discrepancy), iScore placed the response into an arbitration queue for review and rescoring by the SR. Readers were aware neither of the discrepancy arbitration nor which of their scored responses was to be arbitrated. Likewise, the SR was not aware of either readers’ scores when iScore presented the response for review. It was only after the SR had applied their own score to the response that the readers’ scores were then made viewable. This SR review and rescoring of the response became the score of record, thereby overriding the readers’ scores.

Readers with a disproportionate number of adjacent and/or discrepant scores in double-blind scoring were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds. Table 4-9 displays the final summary statistics for double-blind scoring.

Table 4-9. 2010–11 Montana CRT Double-Blind Summary Statistics

<i>Content area</i>	<i>Grade</i>	<i>Responses</i>		
		<i>Number scored</i>	<i>Number double-blind scored</i>	<i>Percentage double-blind scored</i>
Mathematics	3	54,431	1,844	3.4
	4	55,036	1,710	3.1
	5	53,911	1,756	3.3
	6	54,660	1792	3.3
	7	54,400	2,062	3.8
	8	54,670	2,482	4.5
	10	53,186	3,892	7.3
Reading	3	22,100	881	4.0
	4	22,260	724	3.3
	5	21,810	603	2.8
	6	22,110	635	2.9
	7	22,114	807	3.6
	8	21,116	837	4.0
	10	21,520	1,071	5.0
Science	4	22,260	648	2.9
	8	21,116	959	4.5
	10	21,520	1,182	5.5

In the case that the individual was not allowed to resume scoring, however, the content area chief reader had the right to remove (“void”) all of that reader’s scores applied to the item from that day’s work totals. Voided responses in iScore were returned to the response queue and rescored by readers able to maintain the scoring accuracy standard. Table 4-10 summarizes the statistics relevant to voided or blocked readers.

Table 4-10. 2010–11 Montana CRT Voided or Blocked Reader Statistics

<i>Content area</i>	<i>Grade</i>	<i>Item</i>	<i>Number of readers</i>	
			<i>with voided scores</i>	<i>NOT allowed to continue scoring *</i>
Mathematics	3	23	2	0
		24	0	0
		25	7	0
		48	0	0
		72	3	0
	4	23	0	0
		24	0	0
		25	5	0
		48	0	0
		72	0	0
Mathematics	5	23	0	0
		24	4	0
		25	4	0
		48	0	0
		72	0	0
	6	18	0	0
		19	0	0

continued

Content area	Grade	Item	Number of readers	
			with voided scores	NOT allowed to continue scoring *
Mathematics	6	20	0	0
		23	0	0
		73	4	0
	7	18	0	0
		19	0	0
		20	0	0
		23	1	0
		73	1	0
	8	18	0	0
		19	0	0
		20	0	0
		23	6	0
		73	2	0
	10	18	0	0
		19	0	0
		20	0	0
		23	4	0
		73	3	0
Reading	3	27	2	0
		81	1	0
	4	27	0	0
		81	0	0
	5	27	0	0
		81	0	0
Reading	6	27	0	0
		81	0	0
	7	27	2	0
		81	4	0
	8	27	7	0
		81	1	0
Science	10	27	0	0
		81	0	0
	4	23	0	0
		69	0	0
	8	23	0	0
		69	0	0
	10	23	3	0
		69	3	0

* Based upon other quality monitoring (read-behinds and double-blinds)

CHAPTER 5. CLASSICAL ITEM ANALYSIS

As noted in Brown (1983), “A test is only as good as the items it contains.” A complete evaluation of a test’s quality must include an evaluation of each item. Both *Standards for Educational and Psychological Testing* (AERA et al., 1999) and *Code of Fair Testing Practices in Education* (2004) include standards for identifying quality items. Items should assess only knowledge or skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students in particular racial, ethnic, or gender groups.

Both qualitative and quantitative analyses are conducted to ensure that Montana CRT items meet these standards. Qualitative analyses are described in earlier chapters of this report; this chapter focuses on quantitative evaluations. Statistical evaluations are presented in four parts: 1) difficulty indices, 2) item-test correlations, 3) differential item functioning (DIF) statistics, and 4) dimensionality analyses. The item analyses presented here are based on the statewide administration of the Montana CRT in spring 2011. Note that the information presented in this chapter is based on the items common to all forms, since those are the items on which student scores are calculated. (Item analyses are also performed for field-test items, and the statistics are then used during the item review process and form assembly for future administrations.)

5.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All multiple-choice, constructed-response, and short-answer items are evaluated in terms of item difficulty according to standard classical test theory practices. Difficulty is defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item. Multiple-choice and short-answer items are scored dichotomously (correct vs. incorrect), so for these items the difficulty index is simply the proportion of students who correctly answered the item. Constructed-response items are scored polytomously, meaning that a student can achieve a score of 0, 1, 2, 3, or 4. By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale, ranging from 0.0 to 1.0 regardless of the item type. Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an *easiness* index, because larger values indicate easier items. An index of 0.0 indicates that all students received no credit for the item, and an index of 1.0 indicates that all students received full credit for the item.

Items that are answered correctly by almost all students provide little information about differences in student abilities, but they do indicate knowledge or skills that have been mastered by most students. Similarly, items that are correctly answered by very few students provide little information about differences in student abilities, but may indicate knowledge or skills that have not yet been mastered by most students. In general, to provide the best measurement, difficulty indices should range from near-chance performance (0.25 for four-

option multiple-choice items or essentially zero for constructed-response or short-answer items) to 0.90, with the majority of items generally falling between around 0.4 and 0.7. However, on a standards-referenced assessment such as the Montana CRT, it may be appropriate to include some items with very low or very high item difficulty values to ensure sufficient content coverage.

A desirable characteristic of an item is for higher-ability students to perform better on the item than lower-ability students do. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of the item. Within classical test theory, the item-test correlation is referred to as the item's discrimination, because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. For constructed-response items, the item discrimination index used was the Pearson product-moment correlation; for dichotomous items (multiple-choice and short-answer), the corresponding statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is -1.0 to 1.0, with a typical observed range from 0.2 to 0.6.

Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the discrimination index can be thought of as a measure of construct consistency.

A summary of the item difficulty and item discrimination statistics for each subject/grade combination is presented in Table 5-1. Note that the statistics are presented for all items as well as by item type (multiple-choice and open-response, which includes both constructed-response and short-answer items). The mean difficulty and discrimination values shown in the table are within generally acceptable and expected ranges.

Table 5-1. 2010–11 Montana CRT: Summary of Item Difficulty and Discrimination Statistics by Subject and Grade

Subject	Grade	Item type	Number of items	p-value		Discrimination	
				Mean	Standard deviation	Mean	Standard deviation
Mathematics	3	All	60	0.69	0.18	0.37	0.08
		MC	55	0.69	0.18	0.37	0.08
		OR	5	0.60	0.16	0.42	0.10
	4	All	60	0.67	0.16	0.37	0.08
		MC	55	0.67	0.16	0.36	0.07
		OR	5	0.59	0.13	0.48	0.08
	5	All	60	0.63	0.15	0.38	0.10
		MC	55	0.63	0.15	0.37	0.10
		OR	5	0.63	0.11	0.47	0.10
	6	All	60	0.59	0.14	0.36	0.11
		MC	55	0.59	0.14	0.35	0.11
		OR	5	0.54	0.21	0.47	0.12
	7	All	60	0.59	0.16	0.39	0.09
		MC	55	0.60	0.17	0.37	0.08
		OR	5	0.53	0.13	0.53	0.05
	8	All	60	0.57	0.15	0.36	0.09
		MC	55	0.58	0.16	0.35	0.07

continued

Subject	Grade	Item type	Number of items	p-value		Discrimination	
				Mean	Standard deviation	Mean	Standard deviation
Mathematics	8	OR	5	0.50	0.06	0.52	0.11
	10	All	60	0.49	0.16	0.36	0.13
		MC	55	0.50	0.15	0.34	0.12
		OR	5	0.39	0.18	0.57	0.09
Reading	3	All	54	0.71	0.15	0.39	0.08
		MC	52	0.72	0.14	0.39	0.07
		OR	2	0.38	0.04	0.52	0.04
	4	All	54	0.71	0.13	0.38	0.07
		MC	52	0.73	0.12	0.38	0.07
		OR	2	0.40	0.08	0.44	0.01
	5	All	54	0.73	0.12	0.38	0.08
		MC	52	0.74	0.11	0.37	0.08
		OR	2	0.49	0.02	0.42	0.04
	6	All	54	0.70	0.14	0.38	0.07
		MC	52	0.71	0.13	0.37	0.07
		OR	2	0.43	0.03	0.48	0.00
	7	All	54	0.74	0.10	0.42	0.07
		MC	52	0.75	0.09	0.42	0.06
		OR	2	0.51	0.03	0.58	0.04
	8	All	54	0.72	0.09	0.42	0.07
		MC	52	0.73	0.09	0.41	0.07
		OR	2	0.59	0.01	0.59	0.01
	10	All	54	0.72	0.11	0.39	0.07
		MC	52	0.73	0.10	0.38	0.07
		OR	2	0.49	0.03	0.57	0.07
Science	4	All	55	0.67	0.15	0.32	0.08
		MC	53	0.67	0.15	0.31	0.08
		OR	2	0.54	0.10	0.45	0.06
	8	All	55	0.60	0.17	0.32	0.10
		MC	53	0.61	0.17	0.32	0.09
		OR	2	0.31	0.11	0.58	0.04
	10	All	55	0.58	0.14	0.34	0.10
		MC	53	0.59	0.13	0.33	0.09
		OR	2	0.36	0.05	0.51	0.07

MC = multiple-choice; OR = open-response

A comparison of indices across grade levels is complicated because these indices are population dependent. Direct comparisons would require that either the items or students were common across groups. Since that is not the case, it cannot be determined whether differences in performance across grade levels are due to differences in student abilities, differences in item difficulties, or both. With this caveat in mind, it appears generally that for mathematics and, to a lesser extent, science, students in higher grades found their items more difficult than students in lower grades found theirs, while in reading, the difficulty values are fairly constant across grades.

Comparing the difficulty indices of multiple-choice items and constructed-response or short-answer items is inappropriate because multiple-choice items can be answered correctly by guessing. Thus, it is not surprising that the difficulty indices for multiple-choice items tend to be higher (indicating that students performed better on these items) than the difficulty indices for constructed-response items. Similarly,

discrimination indices for the four-point constructed-response items were larger than those for the dichotomous items due to the greater variability of the former (i.e., the partial credit these items allow) and the tendency for correlation coefficients to be higher given greater variances of the correlates.

In addition to the item difficulty and discrimination summaries presented above, item level classical statistics and item level score distributions were also calculated. Item level classical statistics are provided in Appendix E; item difficulty and discrimination values are presented for each item. The item difficulty and discrimination indices are within generally acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall. There were a small number of items with near-zero discrimination indices, but none were negative. While it is not inappropriate to include items with low discrimination values or with very high or very low item difficulty values to ensure that content is appropriately covered, there were very few such cases on the Montana CRT. Item level score-point distributions are provided for constructed-response items in Appendix F; for each item, the percentage of students who received each score point is presented.

5.2 DIFFERENTIAL ITEM FUNCTIONING

Code of Fair Testing Practices in Education (2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit and that actions should be taken to ensure that differences in performance are due to construct-relevant, rather than irrelevant, factors. *Standards for Educational and Psychological Testing* (AERA et al., 1999) includes similar guidelines. As part of the effort to identify such problems, Montana CRT items were evaluated in terms of differential item functioning (DIF) statistics.

For the Montana CRT, the standardization DIF procedure (Dorans and Kulick, 1986) was employed to evaluate subgroup differences. The standardization DIF procedure is designed to identify items for which subgroups of interest perform differently, beyond the impact of differences in overall achievement. The DIF procedure calculates the difference in item performance for two groups of students (at a time) matched for achievement on the total test. Specifically, average item performance is calculated for students at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups.

Computed DIF indices have a theoretical range from -1.0 to 1.0 for multiple-choice and short-answer items, and the index is adjusted to the same scale for constructed-response items. Dorans and Holland (1993) suggested that index values between -0.05 and 0.05 should be considered negligible. Dorans and Holland further stated that items with values between -0.10 and -0.05 and between 0.05 and 0.10 (i.e., “low” DIF)

should be inspected to ensure that no possible effect is overlooked, and that items with values outside the -0.10 to 0.10 range (i.e., “high” DIF) are more unusual and should be examined very carefully.¹

When differential performance between two groups occurs on an item (i.e., a DIF index in the “low” or “high” categories), it may or may not be indicative of item bias. Course-taking patterns or differences in school curricula can lead to DIF but for construct-relevant reasons. On the other hand, if subgroup differences in performance could be traced to differential experience (such as geographical living conditions or access to technology), the inclusion of such items should be reconsidered.

For the 2010–11 Montana CRT, six subgroup comparisons were evaluated for DIF:

- Male versus female
- White versus Native American
- White versus Hispanic
- Disability versus no disability
- Low income versus not low income
- Limited English proficient versus Not limited English proficient

Other comparisons (e.g., other ethnic group pairs) were not analyzed using DIF procedures, because limited sample sizes would have inflated type I error rates. The tables in Appendix G present the number of items classified as either “low” or “high” DIF, overall and by group favored.

5.3 DIMENSIONALITY ANALYSIS

The DIF analyses outlined in the previous section were performed to identify items that showed evidence of differences in performance between pairs of subgroups beyond that which would be expected based on the primary construct that underlies total test score (also known as the “primary dimension”; for example, general achievement in math). When items are flagged for DIF, statistical evidence points to their measuring an additional dimension(s) to the primary dimension.

Because tests are constructed with multiple content area subcategories, and their associated knowledge and skills, the potential exists for a large number of dimensions being invoked beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, the primary dimension they share typically explains an overwhelming majority of variance in test scores. In fact, the presence of just such a dominant primary dimension is the psychometric assumption that provides the foundation for the unidimensional item response theory (IRT) models that are used for calibrating, linking,

¹ It should be pointed out here that DIF for items is evaluated initially at the time of field testing. If an item displays high DIF, it is flagged for review by a Measured Progress content specialist. The content specialist consults with the Department of Education to determine whether to include the flagged item in a future operational test administration.

scaling, and equating the 2010–11 MontCAS test forms. As noted in the previous section, a statistically significant DIF result does not automatically imply that an item is measuring an irrelevant construct or dimension. An item could be flagged for DIF because it measures one of the construct-relevant dimensions of a subcategory’s knowledge and skills.

The purpose of dimensionality analysis is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (a) the degree to which unidimensionality is violated and (b) the nature of the multidimensionality. Findings from dimensionality analyses performed on the 2010–11 MontCAS common items for mathematics, reading, and science are reported below. (Note: only common items were analyzed since only they are used for score reporting.)

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Both of these methods use as their basic statistical building block the estimated average conditional covariances for item pairs. A conditional covariance is the covariance between two items conditioned on expected total score for the rest of the test, and the average conditional covariance is obtained by averaging over all possible conditioning scores. When a test is strictly unidimensional, all conditional covariances are expected to take on values within random noise of zero, indicating statistically independent item responses for examinees with equal expected total test scores. Nonzero conditional covariances are essentially violations of the principle of local independence, and local dependence implies multidimensionality. Thus, nonrandom patterns of positive and negative conditional covariances are indicative of multidimensionality.

DIMTEST is a hypothesis-testing procedure for detecting violations of local independence. The data are first divided into a training sample and a cross-validation sample. Then an exploratory analysis of the conditional covariances is conducted on the training sample data to find the cluster of items that displays the greatest evidence of local dependence. The cross-validation sample is then used to test whether the conditional covariances of the selected cluster of items displays local dependence, conditioning on total score on the non-clustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

DETECT is an effect-size measure of multidimensionality. As with DIMTEST, the data are first divided into a training sample and a cross-validation sample. The training sample is used to find a set of mutually exclusive and collectively exhaustive clusters of items that best fit a systematic pattern of positive conditional covariances for pairs of items from the same cluster and negative conditional covariances from different clusters. Next, the clusters from the training sample are used with the cross-validation sample data to average the conditional covariances: within-cluster conditional covariances are summed, from this sum the between-cluster conditional covariances are subtracted, this difference is divided by the total number of item pairs, and this average is multiplied by 100 to yield an index of the average violation of local independence for an item pair. DETECT values less than 0.2 indicate very weak multidimensionality (or near

unidimensionality); values of 0.2 to 0.4, weak to moderate multidimensionality; values of 0.4 to 1.0, moderate to strong multidimensionality; and values greater than 1.0, very strong multidimensionality.

DIMTEST and DETECT were applied to the 2010–11 MontCAS. The data for each grade and content area were split into a training sample and a cross-validation sample. Every grade/content area combination had at least 10,000 student examinees, so every training sample and cross-validation sample had at least 5,000 students. DIMTEST was then applied to every grade/content area. DETECT was applied to each dataset for which the DIMTEST null hypothesis was rejected in order to estimate the effect size of the multidimensionality.

Because of the large sample sizes of the Montana tests, DIMTEST would be sensitive even to quite small violations of unidimensionality, and the null hypothesis was rejected at a significance level of 0.01 for every dataset. The rejection of the null hypothesis of unidimensionality for every test was not surprising because strict unidimensionality is an idealization that almost never holds exactly for a given dataset. Thus, it was important to use DETECT to estimate the effect size of the violations of local independence found by DIMTEST. Table 5-2 displays the multidimensional effect size estimates from DETECT.

Table 5-2. MontCAS: Multidimensionality Effect Sizes by Grade and Subject.

<i>Subject</i>	<i>Grade</i>	<i>Multidimensionality effect size</i>	
		<i>2009–10</i>	<i>2010–11</i>
Reading	3	0.10	0.10
	4	0.11	0.12
	5	0.09	0.13
	6	0.09	0.07
	7	0.10	0.09
	8	0.10	0.17
	10	0.14	0.12
	Average	0.10	0.11
Mathematics	3	0.08	0.14
	4	0.13	0.11
	5	0.13	0.14
	6	0.15	0.11
	7	0.14	0.10
	8	0.11	0.11
	10	0.12	0.11
	Average	0.12	0.12
Science	4	0.11	0.11
	8	0.18	0.09
	10	0.12	0.08
	Average	0.14	0.09

All the DETECT values for 2010–11 indicated very weak multidimensionality. The average DETECT values for the three content areas were 0.12 for mathematics, 0.11 for reading, and 0.09 for science. Also shown in Table 5-2 are the values reported in last year’s dimensionality analyses. The DETECT indices for the individual content areas for each grade are seen to be very similar between the two years. In particular,

both sets of values indicate very weak multidimensionality for all the tests, and consequently, the averages for the three content areas for 2009–10 (0.12 for mathematics, 0.10 for reading, and 0.14 for science) are similar to the 2010–11 averages. We also investigated how DETECT divided the tests into clusters to see if there were any discernable patterns with respect to item type—that is, multiple-choice (MC) and constructed-response (CR). Because there were only two constructed-response items at each grade level for each content area, it was difficult to judge whether the clusters produce a significant separation of the multiple-choice and constructed-response items. The strongest separation (and the only clear case of a significant separation) occurred with grade 3 mathematics which had two clusters that together had a total of four items with two of the items being the constructed-response items. The next strongest separations occurred with grades 4 and 5 mathematics, each of which had a single cluster that contained the two constructed-response items along with seven (grade 4) and eight (grade 5) multiple-choice items. In both of these cases, of course, the multiple-choice items clearly outnumbered the constructed-response items. No other cases displayed significant separation of multiple-choice and constructed-response items. This lack of separation of multiple-choice and constructed-response items also occurred in the 2009–10, 2008–09 and 2007–08 tests. A more thorough investigation employing experts in the substantive content of the test forms may result in identification of clusters related to the skills and knowledge areas measured by the items. In any case the violations of local independence from all such effects, as evidenced by the DETECT effect sizes, were very small and do not warrant any changes in test design or scoring.

CHAPTER 6. IRT SCALING AND EQUATING

This chapter describes the procedures used to calibrate, equate, and scale the Montana CRT. During the course of these psychometric analyses, a number of quality control procedures and checks on the processes were implemented. These procedures included evaluation of the calibration processes (e.g., checking the number of Newton cycles required for convergence for reasonableness, checking item parameters and their standard errors for reasonableness, or examining test characteristic curves [TCC] and test information functions [TIF] for reasonableness), evaluation of model fit, evaluation of equating items (e.g., delta analyses, rescore analyses, examination of *b*-plots for reasonableness), and evaluation of the scaling results (e.g., parallel processing by the Psychometrics and Research and Data Analysis departments, comparing lookup tables to the previous year's). An equating report, which provided complete documentation of the quality control procedures and results, was reviewed by the Montana Department of Education (MDE) and approved prior to production of student reports (Measured Progress Department of Psychometrics and Research, *2010–11 MontCAS Criterion-Referenced Test Equating Report*, unpublished manuscript).

Table 6-1 lists items that required intervention either during item calibration or as a result of the evaluations of the equating items. For each flagged item, the table shows the reason it was flagged (e.g., the item was flagged as a result of the delta analyses) and what action was taken. The number of items identified for evaluation was typical across grades and content areas. Descriptions of the evaluations and results are included in the Item Response Theory Results and Equating Results sections below.

Table 6-1. 2010–11 Montana CRT: Items that Required Intervention During IRT Calibration and Equating

<i>IREF</i>	<i>Subject</i>	<i>Grade</i>	<i>Reasons</i>	<i>Action</i>
242748	MAT	03	Delta Analysis	Removed from equating
76891	MAT	04	Delta Analysis	Removed from equating
248007	MAT	04	Delta Analysis	Removed from equating
242883	MAT	05	Delta Analysis	Removed from equating
242999	MAT	05	Delta Analysis	Removed from equating
77531	MAT	06	Delta Analysis	Removed from equating
44181	MAT	08	Delta Analysis	Removed from equating
146573	MAT	10	Delta Analysis	Removed from equating
151783	REA	04	Delta Analysis	Removed from equating
66634	REA	04	Delta Analysis	Removed from equating
150442	REA	05	Delta Analysis	Removed from equating
148854	REA	07	Delta Analysis	Removed from equating
149262	REA	08	Delta Analysis	Removed from equating
56225	SCI	04	Delta Analysis	Removed from equating
89860	SCI	08	Delta Analysis	Removed from equating
75726	SCI	10	Delta Analysis	Removed from equating

6.1 ITEM RESPONSE THEORY

All Montana CRT items were calibrated using item response theory (IRT). IRT uses mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta (θ), and the probability (p) of getting a dichotomous item correct or of getting a particular score on a polytomous item (Hambleton, Swaminathan, & Rogers, 1991; Hambleton & Swaminathan, 1985). In IRT, it is assumed that all items are independent measures of the same construct (i.e., of the same θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and p (Hambleton & van der Linden, 1997; Hambleton & Swaminathan, 1985). The process of determining the specific mathematical relationship between θ and p is called item calibration. After items are calibrated, they are defined by a set of parameters that specify a nonlinear, monotonically increasing relationship between θ and p . Once the item parameters are known, an estimate of θ for each student can be calculated. This estimate, $\hat{\theta}$, is considered to be an estimate of the student's true score or a general representation of student performance. It has characteristics that may be preferable to those of raw scores for equating purposes.

For the 2010–11 CRT, the one-parameter logistic (1PL) model, which can be simplified from the three-parameter logistic (3PL) model, was used for dichotomous items (Hambleton & van der Linden, 1997; Hambleton, Swaminathan, & Rogers, 1991), and the partial credit model (PCM), which can be simplified from the generalized partial credit model, was used for polytomous items (Nering & Ostini, 2010). The 3PL model for dichotomous items can be defined as

$$P_i(1|\theta_j, \xi_i) = c_i + (1 - c_i) \frac{\exp[D a_i (\theta_j - b_i)]}{1 + \exp[D a_i (\theta_j - b_i)]},$$

where
 i indexes the items,
 j indexes students,
 a represents item discrimination,
 b represents item difficulty,
 c is the pseudo guessing parameter,
 ξ_i represents the set of item parameters (a , b , and c), and

D is a normalizing constant equal to 1.701.

In the case of the Montana CRT, the a_i term in the equation is equal to 1.0 and the c_i term is equal to 0.0 for all items, which reduces to the 1PL model:

$$P_i(\theta) = \frac{\exp D(\theta - b_i)}{1 + \exp D(\theta - b_i)}$$

For polytomous items, the generalized partial credit model can be defined as

$$P_{jk}(\theta) = \frac{\exp \sum_{v=0}^k [Da_j(\theta - b_j + d_v)]}{\sum_{c=1}^m \exp \sum_{v=1}^c [Da_j(\theta - b_j + d_v)]},$$

where
 j indexes items,
 k indexes students,
 a represents item discrimination,
 b represents item difficulty,
 d represents category step parameter, and
 D is a normalizing constant equal to 1.701.

In the case of the Montana CRT, the a_j term in the equation is equal to 1.0 for all items.

For more information about item calibration and determination, the reader is referred to Lord and Novick (1968), Hambleton and Swaminathan (1985), or Baker and Kim (2004).

6.2 ITEM RESPONSE THEORY RESULTS

The tables in Appendix H give the IRT item parameters of all common items on the 2010–11 CRT by grade and content area. In addition, Appendix I shows graphs of the test characteristic curves (TCCs) and test information functions (TIFs), which are defined below.

TCCs display the expected (average) raw score associated with each θ_j value between -4.0 and 4.0. Mathematically, the TCC is computed by summing the ICCs of all items that contribute to the raw score. Using the notation introduced in Section 10.1, the expected raw score at a given value of θ_j is

$$E(X | \theta_j) = \sum_{i=1}^n P_i(1 | \theta_j),$$

where
 i indexes the items (and n is the number of items contributing to the raw score),
 j indexes students (here, θ_j runs from -4 to 4), and
 $E(X | \theta_j)$ is the expected raw score for a student of ability θ_j .

The expected raw score monotonically increases with θ_j , consistent with the notion that students of high ability tend to earn higher raw scores than do students of low ability. Most TCCs are “S-shaped”: flatter at the ends of the distribution and steeper in the middle.

The TIF displays the amount of statistical information that the test provides at each value of θ_j . Information functions depict test precision across the entire latent trait continuum. There is an inverse relationship between the information of a test and its standard error of measurement (SEM). For long tests, the SEM at a given θ_j is approximately equal to the inverse of the square root of the statistical information at θ_j (Hambleton, Swaminathan, & Rogers, 1991), as follows:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}$$

Compared to the tails, TIFs are often higher near the middle of the θ distribution where most students are located.

PARSCALE v4.1 (Muraki & Bock, 2003) software was used to perform all IRT analyses for the Montana CRT. Each item occupied only one block in the calibration run, and the 1.701 normalizing constant was used. A default convergence criterion of 0.001 was used. The number of Newton cycles required for convergence for each grade and content area during the IRT analysis can be found in Table 6-2. The number of cycles required fell within acceptable ranges.

Table 6-2. 2010–11 MontCAS CRT: Number of Newton Cycles Required for Convergence

<i>Subject</i>	<i>Grade</i>	<i>Cycles</i>
Mathematics	3	29
	4	20
	5	17
	6	5
	7	9
	8	4
	10	5
Reading	3	35
	4	36
	5	34
	6	32
	7	44
	8	40
	10	37
Science	4	12
	8	5
	10	5

6.3 EQUATING

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year's forms to those given in the previous year. Equating ensures that students are not given an unfair advantage or disadvantage because the test form they took is easier or harder than those taken by other students.

Equating for the Montana CRT used the *anchor-test-nonequivalent-groups* design described by Petersen, Kolen, and Hoover (1989). In this equating design, no assumption is made about the equivalence of the examinee groups taking different test forms (that is, naturally occurring groups are assumed). IRT is particularly useful for equating nonequivalent groups (Allen & Yen, 1979). The fixed common-item IRT procedure was used: The anchor items from the previous year's administration were identified during this year's calibrations, and their IRT parameters were fixed to last year's values. This method results in all person and item parameters being on the same θ scale as they were in the previous year. The procedures used for equating and scaling do not change the ranking of students, give more weight to particular items, or change students' performance level classifications.

6.4 EQUATING RESULTS

An Equating Report was submitted to the OPI for their approval prior to production of student reports. The equating report details the results of a variety of quality control activities that were implemented within the Psychometrics and Research Department during IRT calibration and equating, including examining *b*-plots and TCCs and conducting delta and rescore analyses. The evaluations of the equating results are summarized in Table 6-1 above. The *b*-plots can be found in Appendix J. The procedures used to evaluate equating items are described below.

Appendix K presents the results from the delta analysis. This procedure was used to evaluate the performance of equating items, and the discard status presented in the appendix indicates whether the item was used in equating. As can be seen in the appendix, as well as in Table 6-1, a very small number of items were identified as problematic based on the results of the delta analyses and were excluded from use in equating.

Also presented in Appendix K are the results from the rescore analysis. With this analysis, 200 random papers from the previous year were interspersed with this year's papers to evaluate scorer consistency from one year to the next. All effect sizes were well below the criterion value for excluding an item as an equating item, 0.80 (in absolute value).

6.5 ACHIEVEMENT STANDARDS

Cutpoints for the Montana CRT in reading and mathematics were set at standard-setting meetings held in June and July 2006, and cutpoints in science were set in June 2008. Details of the standard-setting procedures can be found in the standard-setting reports and technical reports of those years. The cuts on the theta scale that were established at those meetings are presented in Table 6-3 below. The θ -metric cut scores that emerged from the standard-setting meetings will remain fixed throughout the assessment program unless standards are reset for any reason. Also shown in the table are the cutpoints on the reporting score scale (described below).

Table 6-3. 2010–11 Montana CRT: Cut Scores on the Theta Metric and Reporting Scale by Subject and Grade

Subject	Grade	Theta			Scaled score				
		Cut 1	Cut 2	Cut 3	Minimum	Cut 1	Cut 2	Cut 3	Maximum
Mathematics	3	-0.54340	-0.20337	0.44500	200	225	250	290	300
	4	-0.29081	0.05530	0.65734	200	225	250	291	300
	5	-0.55315	-0.20313	0.38248	200	225	250	289	300
	6	-0.55054	-0.17902	0.36958	200	225	250	287	300
	7	-0.51684	-0.16514	0.35144	200	225	250	289	300
	8	-0.52251	-0.09914	0.46022	200	225	250	283	300
	10	-0.57541	-0.06623	0.50451	200	225	250	281	300
Reading	3	-1.03019	-0.52098	0.26228	200	225	250	287	300
	4	-0.64979	-0.19215	0.55362	200	225	250	289	300
	5	-0.86117	-0.43483	0.24763	200	225	250	287	300
	6	-0.82220	-0.42340	0.26115	200	225	250	289	300
	7	-0.87767	-0.44082	0.29929	200	225	250	288	300
	8	-0.54622	-0.17634	0.50092	200	225	250	289	300
	10	-0.42862	-0.08340	0.55241	200	225	250	289	300
Science	4	-0.70081	-0.14474	0.55956	200	225	250	282	300
	8	-0.57275	-0.07715	0.58285	200	225	250	283	300
	10	-0.37793	0.12744	0.52244	200	225	250	270	300

6.5.1 Distributions

Table L-1 in Appendix L shows performance level distributions for each of the last three years by subject and grade.

6.6 SCALED SCORES

6.6.1 Description of Scale

Montana CRT scores in each content area are reported on a scale ranging from 200 to 300. By providing information that is more specific about the position of a student's results, scaled scores supplement

performance level scores. School- and district-level scaled scores are calculated by computing the average of student-level scaled scores. Students' raw scores (i.e., total number of points) on the 2010–11 Montana CRT were translated to scaled scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another. In the same way that a given temperature can be expressed on either Fahrenheit or Celsius scales, or the same distance can be expressed in either miles or kilometers, student scores on the 2010–11 Montana CRT tests can be expressed in raw or scaled scores.

It is important to note that converting from raw scores to scaled scores does not change students' performance level classifications. Given the relative simplicity of raw scores, it is fair to ask why scaled scores instead of raw scores are used in Montana CRT reports. Foremost, scaled scores offer the advantage of simplifying result reporting across content areas, grade levels, and subsequent years. Because the standard-setting process typically results in different cut scores across content areas on a raw score basis, it is useful to transform these raw cut scores to a scale that is more easily interpretable and consistent. For the Montana CRT, a score of 225 is the cut score between the Novice and Nearing Proficiency performance levels. This is true regardless of content area, grade level, or year. For example, the raw cut score between Novice and Nearing Proficiency may be 35 in grade 8 mathematics but may be 33 in grade 10 mathematics. Using scaled scores greatly simplifies the task of understanding how a student performed. The raw score to scaled score look-up tables for each content area and grade are presented in Appendix M.

6.6.2 Calculations

For Montana CRT, scaled scores were obtained by a simple translation of students' scores using a linear equation of the form

$$SS = mY + b,$$

where
 m is the slope,
 b is the intercept, and
 Y represents the student's score.

A separate linear transformation was used for each grade/content area combination. Each line was determined by using threshold values obtained via standard setting and fixing the Novice/Nearing Proficiency and Nearing Proficiency/Proficient scaled score cuts to 225 and 250, respectively. The cut between Proficient and Advanced was then allowed to vary across grades and content areas. The scaled score values obtained using this formula were rounded to the nearest integer and truncated, as necessary, so that no student received a score lower than 200 or higher than 300.

For science, the student score used for scaling was the ability estimate on the theta scale, $\hat{\theta}$, which was found from the students' raw scores by mapping through the TCC. For reading and mathematics, on the other hand, scaling was done from raw score. As with science, the students' raw scores on the 2010–11 test were transformed into ability estimates on the theta scale using the TCC. These ability estimates were then

transformed into an expected raw score on the reference test form (2005–06, when standards were established for reading and mathematics) using the TCC for the reference test. This expected raw score was then scaled onto the reporting metric.

Table 6-4 shows the scaling constants by subject and grade.

Table 6-4. 2010–11 Montana CRT: Scaled Score Slope and Intercept by Subject and Grade

<i>Subject</i>	<i>Grade</i>	<i>Slope</i>	<i>Intercept</i>
Mathematics	3	3.1692	118.5242
	4	3.0431	141.4551
	5	2.8083	155.7965
	6	2.7906	159.5450
	7	3.0378	159.7850
	8	2.4365	172.4985
	10	2.0947	181.1735
Reading	3	2.4370	182.0623
	4	2.5939	174.3429
	5	2.7798	161.4892
	6	3.0026	154.7492
	7	2.5872	169.9388
	8	3.0898	145.1710
	10	3.1680	130.2323
Science	4	44.9584	256.5073
	8	50.4439	253.8917
	10	49.4687	243.6957

6.6.3 Distributions

Graphs of the scaled score cumulative frequency distributions for the last three years are presented in Appendix L. Note that the graphs show the percent of students at or below each scaled score, thus the lowest line in a given graph depicts the highest performing group. For example, in the graph for grade 5 mathematics (Figure L-3), the line showing the cumulative distribution for 2010–11 is consistently lower than that for 2009–10 which, in turn, is consistently lower than that for 2008–09. This pattern indicates that student performance on the grade 5 mathematics test has improved in each of the last two years.

CHAPTER 7. RELIABILITY

Although an individual item's performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way items function together and complement one another. Tests that function well provide a dependable assessment of the student's level of ability. Unfortunately, no test can do this perfectly. A variety of factors can contribute to a given student's score being either higher or lower than his or her true ability. For example, a student may misread an item, or mistakenly fill in the wrong bubble when he or she knew the answer. Collectively, extraneous factors that impact a student's score are referred to as measurement error. Any assessment includes some amount of measurement error; that is, no measurement is perfect. This is true of all academic assessments—some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores or vice versa. Consequently, one cannot reliably measure a student's true level of ability with such a test. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as reliable.

There are a number of ways to estimate an assessment's reliability. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, then the extraneous factors affecting performance are small and the test is reliable. (This is referred to as "test-retest reliability.") A potential problem with this approach is that students may remember items from the first administration or may have gained (or lost) knowledge or skills in the interim between the two administrations. A solution to the "remembering items" problem is to give a different, but parallel test at the second administration. If student scores on each test correlate highly the test is considered reliable. (This is known as "alternate forms reliability," because an alternate form of the test is used in each administration.) This approach, however, does not address the problem that students may have gained (or lost) knowledge or skills in the interim between the two administrations. In addition, the practical challenges of developing and administering parallel forms generally preclude the use of parallel forms reliability indices. One way to address the latter problems is to split the test in half and then correlate students' scores on the two half-tests; this in effect treats each half-test as a complete test. By doing this, the problems associated with an intervening time interval and of creating and administering two parallel forms of the test are alleviated. This is known as a "split-half estimate of reliability." If the two half-test scores correlate highly, items on the two half-tests must be measuring very similar knowledge or skills. This is evidence that the items complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test halves will result in a different correlation. Another problem with the split-half method of calculating reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter

test is less reliable than a longer test. Cronbach (1951) provided a statistic, α (alpha), which eliminates the problem of the split-half method by comparing individual item variances to total test variance. Cronbach's α was used to assess the reliability of the 2010–11 Montana CRT:

$$\alpha \equiv \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n \sigma^2_{(Y_i)}}{\sigma_x^2} \right],$$

where
 i indexes the item,
 n is the total number of items,
 $\sigma^2_{(Y_i)}$ represents individual item variance, and
 σ_x^2 represents the total test variance.

7.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Table 7-1 presents descriptive statistics, Cronbach's α coefficient, and raw score standard errors of measurement (SEMs) for each content area and grade. (Statistics are based on common items only.)

Table 7-1. 2010–11 Montana CRT: Raw Score Descriptive Statistics, Cronbach's Alpha and Standard Errors of Measurement (SEM) by Subject and Grade

Subject	Grade	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
Mathematics	3	10,492	66	44.04	11.54	0.91	3.53
	4	10,544	66	43.12	11.74	0.91	3.57
	5	10,366	66	41.10	12.45	0.91	3.67
	6	10,522	66	37.29	12.08	0.91	3.67
	7	10,482	66	38.38	12.44	0.92	3.55
	8	10,448	66	37.19	12.32	0.91	3.75
	10	10,059	66	31.36	12.36	0.91	3.73
Reading	3	10,462	60	40.56	10.35	0.91	3.11
	4	10,524	60	40.89	9.99	0.91	3.03
	5	10,356	60	42.04	9.93	0.90	3.11
	6	10,523	60	40.50	10.05	0.90	3.12
	7	10,495	60	43.20	11.08	0.92	3.08
	8	10,484	60	42.50	11.45	0.92	3.19
	10	10,077	60	41.87	10.69	0.91	3.18
Science	4	10,574	61	39.79	9.63	0.87	3.44
	8	10,483	61	34.89	9.97	0.88	3.48
	10	10,077	61	34.12	10.37	0.89	3.49

For mathematics, the reliability coefficients ranged from 0.91 to 0.92; for reading, from 0.90 to 0.92; and for science, from 0.87 to 0.89. Because different grades and content areas have different test designs (e.g.,

the number of items varies by test), it is inappropriate to make inferences about the quality of one test by comparing its reliability to that of another test from a different grade and/or content area.

7.2 2010–11 SUBGROUP RELIABILITY

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2010–11 Montana CRT. Appendix N presents reliabilities for various subgroups of interest. Subgroup Cronbach's α 's were calculated using the formula defined above based only on the members of the subgroup in question in the computations; values are only calculated for subgroups with 10 or more students. For mathematics, subgroup reliabilities ranged from 0.79 to 0.93; for reading, from 0.82 to 0.93; and for science, from 0.74 to 0.90.

For several reasons, the results of this section should be interpreted with caution. First, inherent differences between grades and content areas preclude making valid inferences about the quality of a test based on statistical comparisons with other tests. Second, reliabilities are dependent not only on the measurement properties of a test but on the statistical distribution of the studied subgroup. For example, it can be readily seen in Appendix N that subgroup sample sizes may vary considerably, which results in natural variation in reliability coefficients. Additionally, α , which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998). Third, there is no industry standard to interpret the strength of a reliability coefficient, and this is particularly true when the population of interest is a single subgroup.

7.3 REPORTING SUBCATEGORY RELIABILITY

Of even more interest are reliabilities for the reporting subcategories within Montana CRT content areas, described in Chapter 3. Cronbach's α coefficients for subcategories were calculated via the same formula defined previously using just the items of a given subcategory in the computations. Results are presented in Appendix N. Once again, as expected, because they are based on a subset of items rather than the full test, computed subcategory reliabilities were lower (sometimes substantially so) than were overall test reliabilities, and interpretations should take this into account.

For mathematics, subcategory reliabilities ranged from 0.43 to 0.85; for reading, from 0.55 to 0.81; and for science, from 0.18 to 0.70. The subcategory reliabilities were lower than those based on the total test and approximately to the degree one would expect based on classical test theory. Qualitative differences between grades and content areas once again preclude valid inferences about the quality of the full test based on statistical comparisons among subtests.

7.4 INTERRATER CONSISTENCY

Chapter 4 of this report describes in detail the processes that were implemented to monitor the quality of the hand-scoring of student responses for short-answer and constructed-response items. One of these

processes was double-blind scoring: approximately 2% of student responses were randomly selected and scored independently by two different scorers. Results of the double-blind scoring were used during scoring to identify scorers who required retraining or other intervention and are presented here as evidence of the reliability of the Montana CRT. A summary of the interrater consistency results is presented in Table 7-2 below. Results in the table are collapsed across the hand-scored items by subject, grade, and number of score categories (two for short-answer items and five for constructed-response items). The table shows the number of included scores, the percent exact agreement, the percent adjacent agreement, the correlation between the first two sets of scores, and the percent of responses that required a third score. This same information is provided at the item level in Appendix O.

Table 7-2. 2010–11 Montana CRT: Summary of Interrater Consistency Statistics Collapsed Across Items by Subject and Grade

<i>Subject</i>	<i>Grade</i>	<i>Number of score categories</i>	<i>Number of included scores</i>	<i>Percent exact</i>	<i>Percent adjacent</i>	<i>Correlation</i>	<i>Percent of third scores</i>
Mathematics	3	2	654	98.62	1.38	0.97	0.00
		5	494	80.77	17.00	0.93	2.23
	4	2	642	98.44	1.56	0.97	0.00
		5	463	84.02	12.31	0.91	3.67
	5	2	821	98.17	1.83	0.96	0.00
		5	450	79.11	18.89	0.92	2.00
	6	2	643	98.91	1.09	0.97	0.00
		5	450	86.67	12.00	0.94	1.33
	7	2	635	98.74	1.26	0.97	0.00
		5	427	83.14	15.93	0.93	0.94
	8	2	631	98.10	1.90	0.96	0.00
		5	458	84.93	13.76	0.95	1.09
	10	2	584	98.97	1.03	0.98	0.00
		5	402	79.60	17.91	0.91	2.49
Reading	3	5	445	63.82	35.51	0.81	0.67
	4	5	431	66.82	30.86	0.71	2.32
	5	5	418	61.72	35.89	0.74	2.39
	6	5	427	67.68	31.62	0.78	0.70
	7	5	448	60.71	38.17	0.81	0.89
	8	5	470	59.15	36.81	0.77	4.04
	10	5	391	58.82	40.15	0.81	1.02
Science	4	5	427	72.13	23.65	0.86	4.22
	8	5	418	83.25	16.27	0.93	0.48
	10	5	396	63.13	32.07	0.75	4.80

7.5 RELIABILITY OF PERFORMANCE LEVEL CATEGORIZATION

While related to reliability, the accuracy and consistency of classifying students into performance categories are even more important statistics in a standards based reporting framework (Livingston & Lewis, 1995). After the performance levels were specified and students were classified into those levels, empirical

analyses were conducted to determine the statistical accuracy and consistency of the classifications. For the Montana CRT, students are classified into one of four performance levels: Novice (N), Nearing Proficiency (NP), Proficient (P), or Advanced (A). This section of the report explains the methodologies used to assess the reliability of classification decisions, and results are given.

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated because errorless test scores do not exist. Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are given to the same group of students. In operational test programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and consistency of classification decisions based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2010–11 Montana CRT because it is easily adaptable to all types of testing formats, including mixed format tests.

The accuracy and consistency estimates reported in Appendix P make use of “true scores” in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. Of course, true scores cannot be observed and so must be estimated. In the Livingston and Lewis method, estimated true scores are used to categorize students into their “true” classifications.

For the 2010–11 Montana CRT, after various technical adjustments (described in Livingston & Lewis, 1995), a four by four contingency table of accuracy was created for each content area and grade, where cell $[i, j]$ represented the estimated proportion of students whose true score fell into classification i (where $i = 1$ to 4) and whose observed score fell into classification j (where $j = 1$ to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following statistical adjustments per Livingston and Lewis (1995), a new four by four contingency table was created for each content area and grade and populated by the proportion of students who would be categorized into each combination of classifications according to the two (hypothetical) parallel test forms. Cell $[i, j]$ of this table represented the estimated proportion of students whose observed score on the first form would fall into classification i (where $i = 1$ to 4) and whose observed score on the second form would fall into classification j (where $j = 1$ to 4). The sum of the diagonal entries (i.e., the proportion of students categorized by the two forms into exactly the same classification) signified overall consistency.

Another way to measure consistency is to use Cohen’s (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. It is calculated using the following formula:

$$\kappa = \frac{(\text{Observed agreement}) - (\text{Chance agreement})}{1 - (\text{Chance agreement})} = \frac{\sum_i C_{ii} - \sum_i C_{i.}C_{.i}}{1 - \sum_i C_{i.}C_{.i}},$$

where

$C_{i.}$ is the proportion of students whose observed performance level would be Level i (where $i = 1-4$) on the first hypothetical parallel form of the test;

$C_{.i}$ is the proportion of students whose observed performance level would be Level i (where $i = 1-4$) on the second hypothetical parallel form of the test;

C_{ii} is the proportion of students whose observed performance level would be Level i (where $i = 1-4$) on both hypothetical parallel forms of the test.

Because κ is corrected for chance, its values are lower than are other consistency estimates.

7.5.1 Decision Accuracy and Consistency Results

The decision accuracy and consistency analyses described above are provided in Table P-1 of Appendix P. The table includes overall accuracy and consistency indices, including kappa. Accuracy and consistency values conditional upon performance level are also given. For these calculations, the denominator is the proportion of students associated with a given performance level. For example, the conditional accuracy value is 0.82 for Novice for mathematics grade 3. This figure indicates that among the students whose true scores placed them in this classification, 82 percent would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.74 indicates that 74 percent of students with observed scores in the Novice level would be expected to score in this classification again if a second, parallel test form were used.

For some testing situations, the greatest concern may be decisions around level thresholds. For example, in testing done for NCLB accountability purposes, the primary concern is distinguishing between students who are proficient and those who are not yet proficient. For the 2010–11 Montana CRT, Table P-2 in Appendix P provides accuracy and consistency estimates at each cutpoint as well as false positive and false negative decision rates. (A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.)

The above indices are derived from Livingston and Lewis's (1995) method of estimating the accuracy and consistency of classifications. It should be noted that Livingston and Lewis discuss two versions of the accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An "adjusted" version adjusts the results of one form to match the observed score distribution obtained in the data. The tables use the standard version for two reasons: (1) this "unadjusted" version can be considered a smoothing of the data, thereby decreasing the variability of the results; and (2) for results dealing with the consistency of two parallel forms, the unadjusted tables are symmetrical, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms

that are parallel; that is, it is more intuitive and interpretable for two parallel forms to have the same statistical distribution.

Descriptive statistics relating to the decision accuracy and consistency (DAC) of the 2010–11 Montana CRT tests can be derived from Table P-1. For mathematics, overall accuracy ranged from 0.77 to 0.79, overall consistency ranged from 0.69 to 0.72, and the kappa statistic ranged from 0.56 to 0.60. For reading, overall accuracy ranged from 0.83 to 0.86, overall consistency ranged from 0.77 to 0.81, and the kappa statistic ranged from 0.62 to 0.67. Finally, for science, overall accuracy ranged from 0.75 to 0.78, overall consistency ranged from 0.66 to 0.69, and the kappa statistic was 0.54 for all grade levels. Note that, as with other methods of evaluating reliability, DAC statistics calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Appendix P should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare DAC statistics between grades and content areas.

CHAPTER 8. SCORE REPORTING

The Montana CRT is designed to measure student performance against Montana's content standards. Consistent with this purpose, results on the CRT were reported in terms of performance levels that describe student performance in relation to these established state standards. There are four performance levels: Novice, Nearing Proficiency, Proficient, and Advanced. (Performance level distributions are given in Appendix L.) Students receive a separate performance level classification (based on total scaled score) in each content area.

State results were provided to the OPI via a secure Web site. Reading, mathematics, and science reporting data for the 2010–11 Montana CRT were made available to systems and schools online via the Montana Analysis and Reporting System (MARS) on May 28, 2011. Student reports were delivered to parents on September 20, 2011. Student Reports were also posted online to be accessible to schools. System test coordinators and teachers were also provided with copies of the *Guide to the 2011 Criterion-Referenced Test and CRT-Alternate Assessment Reports* to assist them in understanding the connection between the assessment and the classroom. The guide provides information about the assessment and the use of assessment results.

School- and system-level results are reported as the number and percentages of students attaining each performance level at each grade level tested. As described below, decision rules were formulated in early 2011 by the OPI and Measured Progress to identify students who, during the reporting process, were to be excluded from school- and system-level reports. (A copy of these decision rules is included in this report as Appendix A.) State-level summary reports were also produced.

The reports described in the sections that follow are separated into two categories. The first set of reports described is static reports, which are provided online as PDF documents; student reports are also provided on paper. The static reports are the following:

- Student Report (paper and online)
- School, System, and State Summary Reports (online)

The remaining reports are interactive reports, provided via MARS (see sections 9.3 and 9.4 below):

- Class Roster and Item Level Reports
- Performance Level Summary
- Released Items Summary Data
- Longitudinal Data Report

Sample Report Shells are included as Appendix Q.

8.1 DECISION RULES

As mentioned above, to ensure that reported results for the 2010–11 Montana CRT are accurate relative to collected data and other pertinent information, a document that delineates analysis and reporting rules was created. These decision rules were observed in the analyses of Montana CRT test data and in reporting the test results. Moreover, these rules are the main reference for quality assurance checks.

The decision rules document used for reporting results of the 2011 administration of the Montana CRT is found in Appendix A.

The rules primarily describe the inclusion/exclusion of students at the school-, system-, and state-levels of aggregation. The document also describes rules as they pertain to individual reports. Finally, it describes the classification of students based on their school type or other information provided by the state through the student demographic file (AIM) or collected on the answer booklet.

8.2 STATIC REPORTS

8.2.1 Student Report

The student report is produced for each parent of a student who took or was eligible to take the Montana CRT. The report is shipped to systems and posted online for school/system access.

The student report gives the results for each content area tested. At grades 3, 5, 6, and 7, the content areas are reading and mathematics. At grades 4, 8, and 10, the content areas are reading, mathematics, and science. The student reports give the earned performance level and scaled score for each subject. The report also provides a comparison of the student's performance to that of the state as a whole for each subject. The report contains the results for each subject at the content standard level. The number of points earned by the student in each content standard is reported as well as the range of points earned by students who achieve proficiency.

8.2.2 Summary Reports

The summary report is produced at the school, system, and state levels. The report is produced for each content area in the grade level. For grades 3, 5, 6, and 7, the content areas are reading and mathematics. For grades 4, 8, and 10, the content areas are reading, mathematics, and science. The report consists of three sections: Distribution of Scores, Subtest Results, and Results for Subgroups of Students.

The Distribution of Scores section of the report contains a breakdown of the performance of included students (as described in the decision rules document) into different scaled score intervals. The number and percent of students that fall into each scaled score interval is shown. There is an overall percentage reported for students that fall into each of the four performance levels (Novice, Nearing Proficiency, Proficient, and Advanced). In the *School Summary Report*, the calculations are done at the school, system, and state levels.

The *System Summary Report* contains results at the system and state levels. The *State Summary Report* contains only state level results.

The Subtest Results section of the report summarizes the average points earned in the different content standards by included students (as described in the decision rules document) in the school, system, and state. The average points earned are compared to the total possible points for each content standard.

The Results for Subgroups of Students section of the report summarizes the performance of included students (as described in the decision rules document) broken down by various reporting categories. For each reporting category, the number of tested (included) students is reported as well as the percentage of students in each of the four performance levels. In the *School Summary Report* this is reported at the school, system, and state levels. In the *System Summary Report* the data are reported at the system and state levels. In the *State Summary Report* the data are reported at state level only.

The list of reporting categories is as follows:

- All Students
- Gender
- Ethnicity (American Indian or Alaska native, Asian, Hispanic, Black or African American, Native Hawaiian or Other Pacific Islander, White)
- Special Education
- Students with a 504 Plan
- Title I (optional)
- Tested with Standard Accommodation
- Tested with Nonstandard Accommodation
- Alternate Assessment (results are not given for this category on the Montana CRT Summary reports)
- Migrant
- Gifted/Talented
- LEP/ELL
- Former LEP Student
- LEP Student Enrolled for First Time in a U.S. School
- Free/Reduced Lunch

Data are suppressed if there are less than ten tested (included) in a reporting category at a given aggregation level. New for 2011, data are suppressed in the Distribution of Scores and the Subtest Results sections if there are less than ten tested (included) in a school or system.

The data for the reporting categories were provided by information coded on the students' answer booklets by teachers and/or data supplied by the state through an AIM export. Due to relatively low numbers of students in certain reporting categories, school personnel are advised, under FERPA guidelines, to treat these pages confidentially.

8.3 MONTANA ANALYSIS AND REPORTING SYSTEM (MARS)

Using advanced Web technology, *MARS* gives Montana educators and administrators the ability to filter data based on test year, grade level, content area, standard, and student subgroup. This allows administrators to isolate cross-sections of the results and identify areas of strong or poor performance.

The confidential nature of the data in *MARS* necessitates the strict enforcement of site security. All transmissions are done over Secure Socket Layers (SSL). A system of user role definitions and permissions dictates the scope of access granted to individual users. Organizations (system or school levels) are given administrative power to grant or deny access to their data within the system, and they have the ability to disable users. Personnel using *MARS* may be granted permission to view students' results at an organizational level, or only a select group as defined by the administrator. Predefined reports are included in the system, as is the ability to render and print additional copies.

8.3.1 User Accounts

In *MARS*, principals have the ability to create unique user accounts by assigning specific usernames and passwords to educators in their school such as teachers, curriculum coordinators, or special education coordinators. Once the accounts have been created, individual students may be assigned to each user account. After users have received their usernames and passwords, they are able to log in to their accounts and access the interactive reports which will be populated only with the subgroup of students assigned to them.

Information about the interactive reports and setting up user accounts is available in the *Analysis & Reporting System User Manual* that is available for download on the *MARS* system.

8.4 INTERACTIVE REPORTS

As mentioned above, there are four interactive reports that were available from *MARS*: Roster Report, Performance Level Summary, Released Items Summary Data, and Longitudinal Data. Each of these interactive reports is described in the following sections. Sample interactive reports are provided in Appendix R. To access these four interactive reports, the user clicked the interactive tab on the home page of the system and selected the report desired from the drop down menu. Next, the user applied basic filtering options, such

as the name of the district or school and the grade level/content area test, to open the specific report. At this point, the user had the option of printing the report for the entire grade level or applying advanced filtering options to select a subgroup of students to analyze. Advanced filtering options include gender, ethnicity, limited English proficient (LEP), IEP, low income, migrant, and plan 504. All interactive reports, with the exception of the Longitudinal Data Report, allowed the user to provide a custom title for the report.

8.4.1 Roster Report

The Montana CRT *Roster Report* provides a roster of all students in a school and provides performance on the common items that are released to the public, one report per content area. For all grades and content areas, the student names and identification numbers are listed as row headers down the left side of the report. The items are listed as column headers in the same order they appeared in the released item document.

For each item, the following are shown:

- the depth of knowledge (DOK) code
- the item type
- the correct response key for multiple-choice items
- the total possible points
- content standard

For each student, multiple-choice items are marked either with a plus sign (+), indicating that the student chose the correct multiple-choice response, or a letter (from A to D), indicating the incorrect response chosen by the student. For short-answer and constructed-response items, the number of points earned is shown. All responses to released items are shown in the report, regardless of the student's participation status. The columns on the right side of the report show the Total Test results, broken into several categories. Subcategory Points Earned columns show points earned by the student in each content area subcategory relative to total possible points. A Total Points Earned column is a summary of all points earned and total possible points in the content area. The last two columns show the student's scaled score and performance level. Students reported as Not Tested are given a code in the performance level column to indicate the reason the student did not test. It is important to note that not all items used to compute student scores are included in this report, only released items. At the bottom of the report, the average percentage correct for each multiple-choice item and average scores for the short-answer and constructed-response items are shown for the school, district, and state. When advanced filtering criteria are applied by the user, the School and District Percent Correct/Average Score rows at the bottom of the report are blanked out and only the Group row and the State row for the group selected will contain data. This report can be saved, printed, or exported as a PDF.

The Montana CRT roster is confidential and should be kept secure within the school and district. FERPA requires that access to individual student results be restricted to the student, the student's parents/guardians, and authorized school personnel.

8.4.2 Performance Level Summary

The Performance Level Summary provides a visual display of the percentages of students in each performance level for a selected grade/content area. The four performance levels (Novice, Nearing Proficiency, Proficient, and Advanced) are represented by various colors in a pie chart. A separate table is also included below the chart that shows the number and percentage of students in each performance level. This report can be saved, printed, or exported as a PDF or JPG file.

8.4.3 Item Analysis Data

The Released Items Summary Data report is a school-level report that provides a summary of student responses to the released items for a selected grade/content area. The report is divided into two sections by item type (multiple-choice and open-response). For multiple-choice items, the total number/percent of students who answered the item correctly and the number of students who chose each incorrect option or provided an invalid response are reported. An invalid response on a multiple-choice item is defined as "the item was left blank" or "the student selected more than one option for the item." For open-response items, point value and average score for the item are reported. Users are also able to view the actual released items within this report. If a user clicks on a particular magnifying glass icon next to a released item number, a pop-up box will open displaying the released item.

8.4.4 Longitudinal Data Report

The longitudinal data report is a confidential student-level report that provides individual student performance data for multiple test administrations. Results are reported for a student going back to academic year 2006–07. The state-assigned student identification number is used to link students across test administrations. Student performance on future test administrations will be included on this report over time. This report can be saved, printed, or exported as a PDF file.

8.5 INTERPRETIVE MATERIALS AND WORKSHOPS

An interpretive guide to the reports is provided on the OPI Web site: <http://opi.mt.gov/>.

8.6 QUALITY ASSURANCE

Quality assurance measures are embedded throughout the entire process of analysis and reporting. The data processor, data analyst, and psychometrician assigned to work on Montana CRT implement quality control checks of their respective computer programs and intermediate products. Moreover, when data are

handed off to different functions within the Data Services and Static Reporting (DSSR) and Psychometrics and Research (P&R) departments, the sending function verifies that the data are accurate before handoff. Additionally, when a function receives a data set, the first step is to verify the data for accuracy.

Another type of quality assurance measure is parallel processing. Different exclusions that determine whether each student receives scaled scores and/or is included in different levels of aggregation are parallel processed. Using the decision rules document, two data analysts independently write a computer program that assigns students' exclusions. For each content area and grade combination, the exclusions assigned by each data analyst are compared across all students. Only when 100% agreement is achieved can the rest of data analysis be completed.

Another level of quality assurance involves the procedures implemented by the quality assurance group to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that reported information is correct. The step is conducted in two parts: (1) verify that the computed information was obtained correctly through appropriate application of different decision rules, and (2) verify that the correct data points populate each cell in the Montana CRT reports. The selection of sample schools and districts for this purpose is very specific and can affect the success of the quality control efforts. There are two sets of samples selected that may not be mutually exclusive.

The first set includes those that satisfy the following criteria:

- One-school district
- Two-school district
- Multi-school district

The second set of samples includes districts or schools that have unique reporting situations as indicated by decision rules. This second set is necessary to ensure that each rule is applied correctly. The second set includes those that satisfy the following criteria:

- Private school
- School with excluded (not tested) students

The quality assurance group uses a checklist to implement its procedures. After the checklist is completed, sample reports are circulated for psychometric checks and program management review.

CHAPTER 9. VALIDITY

Because interpretations of test scores, and not a test itself, are evaluated for validity, the purpose of the *2010–11 Montana CRT Technical Report* is to describe several technical aspects of the Montana CRT tests in support of score interpretations (AERA, 1999). Each chapter contributes an important component in the investigation of score validation: test development and design; test administration; scoring, scaling, and equating; item analyses; reliability; and score reporting.

As stated in the overview chapter, *Standards for Educational and Psychological Testing* (AERA et al., 1999) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. The evidence around test content, response processes, internal structure, relationship to other variables, and consequences of testing speak to different *aspects* of validity but are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

Evidence on test content validity is meant to determine how well the assessment tasks represent the curriculum and standards for each content area and grade level. Content validation is informed by the item development process, including how the test blueprints and test items align to the curriculum and standards. Viewed through this lens provided by the Standards, evidence based on test content was extensively described in Chapters 2 and 3. Item alignment with Montana content standards; item bias, sensitivity and content appropriateness review processes; adherence to the test blueprint; use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training are all components of validity evidence based on test content. As discussed earlier, all CRT questions are aligned by Montana educators to specific Montana content standards, and undergo several rounds of review for content fidelity and appropriateness. Items are presented to students in multiple formats (constructed-response, short-answer, and multiple-choice). Finally, tests are administered according to state-mandated standardized procedures, with allowable accommodations, and all test proctors are required to attend annual training sessions.

The scoring information in Chapter 4 describes the steps taken to train and monitor hand-scorers, as well as quality control procedures related to scanning and machine scoring. To speak to student response processes, however, additional studies would be helpful and might include an investigation of students' cognitive methods using think-aloud protocols.

Evidence based on internal structure is presented in great detail in the discussions of item analyses, reliability, and scaling and equating in chapters 5 through 7. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), differential item functioning analyses, dimensionality analyses, reliability, standard errors of measurement, and item response theory parameters and procedures. Each test is equated to the same grade and content area test from the prior year in order to preserve the meaning of scores over time. In general, item difficulty and

discrimination indices were in acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most items were assessing consistent constructs, and students who performed well on individual items tended to perform well overall.

Evidence based on the consequences of testing is addressed in the scaled scores information in Chapter 6 and the reporting information in Chapter 8, as well as in the test interpretation guide, which is a separate document that is referenced in the discussion of reporting. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores. Scaled scores offer the advantage of simplifying the reporting of results across content areas, grade levels, and subsequent years. Performance levels provide users with reference points for mastery at each grade level, which is another useful and simple way to interpret scores. Several different standard reports are provided to stakeholders. In addition, a data analysis tool is provided to each school system to allow educators the flexibility to customize reports for local needs. Additional evidence of the consequences of testing could be supplemented with broader investigation of the impact of testing on student learning.

To further support the validation of the assessment program, additional studies might be considered to provide evidence regarding the relationship of CRT results to other variables including the extent to which scores from the CRT converge with other measures of similar constructs, and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

The evidence presented in this report supports inferences of student achievement on the content represented on the Montana content standards for reading, mathematics, and science for the purposes of program and instructional improvement and as a component of school accountability.

REFERENCES

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Baker, F. B., & Kim, S-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York: Marcel Dekker, Inc.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth: Holt, Rinehart and Winston.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23, 355–368.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description. In P. W. Holland & H. Wainer (Eds.) *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York: John Wiley and Sons, Inc.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer Academic Publishers.
- Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.
- Hambleton, R. K., & van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices. Available from www.apa.org/science/programs/testing/fair-code.aspx
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Muraki, E., & Bock, R. D. (2003). PARSCALE 4.1. Lincolnwood, IL: Scientific Software International.
- Nering, M. L., & Ostini, R. (2010). *Handbook of polytomous item response theory models*. New York: Routledge.

- Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221–262).
- Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, 52, 589–617.
- Stout, W. F., Froelich, A. G., & Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duign, & T. A. B. Snijders (Eds.), *Essays on item response theory*, (pp. 357–375). New York: Springer-Verlag.
- Zhang, J., & Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64, 213–249.

APPENDICES

Appendix A	Analysis and Reporting Decision Rules
Appendix B	Participation Rates
Appendix C	Accommodation Frequencies
Appendix D	Item Review Committee Members
Appendix E	Item Level Classical Statistics
Appendix F	Item Level Score Distributions
Appendix G	Number of Items Classified Into DIF Categories
Appendix H	Item Response Theory Calibration Results
Appendix I	TCCs and TIFs
Appendix J	<i>b</i> -Plots
Appendix K	Analyses of Equating Items (Delta & Rescore Analyses)
Appendix L	Score Distributions
Appendix M	Raw to Scaled Score Look-up Tables
Appendix N	Classical Reliability
Appendix O	Interrater Agreement
Appendix P	Decision Accuracy and Consistency Results
Appendix Q	Report Shells
Appendix R	Sample Interactive Reports

Appendix A—ANALYSIS AND REPORTING DECISION RULES

Analysis and Reporting Decision Rules
Montana Comprehensive Assessment System (MontCAS) CRT and CRT-Alternate
(Final)
Spring 10-11 Administration

This document details rules for analysis and reporting. The final student level data set used for analysis and reporting is described in the “Data Processing Specifications.” This document is considered a draft until the Montana Office of Public Instruction (OPI) signs off. If there are rules that need to be added or modified after said sign-off, OPI sign off will be obtained for each rule. Details of these additions and modifications will be in the Addendum section.

I. General Information

A. Tests Administered

Grade	Subject	Items included in Raw Score		IABS Reporting Categories (Standards) (Not Applicable for CRT-Alternate)
		CRT	CRT-Alt	
03	Reading Math	Common	All	Cat2
04	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3
05	Reading Math	Common	All	Cat2
06	Reading Math	Common	All	Cat2
07	Reading Math	Common	All	Cat2
08	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3
10	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3

B. Reports Produced

1. Student Labels (Printed)
2. Student Report (Printed and posted online)
3. Roster & Item Level Report (CRT-Alt: posted online; CRT:Interactive System)
 - by grade, subject and class/group
4. Summary Report (Online)
 - Consists of sections:

- I. Distribution of Scores
- II. Subtest Results
- III. Results for Subgroups of Students
 - by grade, subject and school
 - by grade, subject and system
 - by grade, subject (state level)

C. Files Produced (Format: comma delimited format)

1. One state file for each grade
 - a. Consists of student level results
 - b. Alternately assessed students are in separate files by grade.
 - c. Naming conventions
 - i. CRT All subjects- Studentdatafile[2 digit grade].csv
 - ii. CRT-Alternate All subjects- altStudentdatafile[2 digit grade].csv
 - d. File layout: Studentdatafilelayout.xls and altstudentdatafilelayout.xls
2. System level files (Format: Excel ; Online)
 - a. Consists of student level results for each system for each grade. Contains all subjects tested at that grade.
 - b. Naming convention: Studentdatafile[2 digit grade].xls
 - c. File Layout: Systemstudentdatafilelayout.xls
3. School level file (Format: Excel; Online)
 - a. Consists of previous year's student level results for each school and grade. Contains all subjects tested at that grade.
 - b. Naming convention: Studentdatafile[2 digit grade].xls
 - c. File Layout: Systemstudentdatafilelayout.xls

D. School Type

Schtype	Source	Description	Included in Aggregations		
			School	System	State
"Pras"	Data file provided by state	Private Accredited School. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No
"Prnas"	Data file provided by state	Private non-accredited school. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No

Schtype	Source	Description	Included in Aggregations		
"SNE"	Scanned data/ updated by OPI	Student not enrolled	No.	No.	No.
"Oth"		Non-private school	Yes	Yes	Yes

E. Other Information

1. CRT are constructed with a combination of common and embedded field test items.
2. The CRT-Alternate consists of a set of 5 performance tasklets. The number of items in each tasklet varies.
3. Braille Students:
 - a. See Appendix A.1 for a list of the items not included in the Braille form.
 - b. If a student is identified as taking the Braille test, these items are not included in the student's raw score. The student is scaled on a separate form based on the items that are available to him or her. See the Calculations section for more information.

II. Student Participation/Exclusions

A. Test Attempt Rules

1. A valid response to a multiple choice item is A, B, C, or D. An asterisk (multiple marks) is not considered a valid response. A valid score for an open response item is a non-blank score.
2. Incomplete (CRT): The student has exactly one (1) valid response to common items.
3. Incomplete (CRT-Alternate): The student has fewer than three (3) scores across all tasklets.
4. The student is classified as Did Not Participate (DNP) in CRT if the student does not have any valid responses for that subject in either CRT or CRT-Alternate and has no not tested reason.

B. Not Tested Reasons

1. If a student is marked First year LEP regardless of items attempted the student is considered first year LEP for reporting purposes. Reading is optional for first year in U.S schools LEP students.

C. Student Participation Status

1. The following students are excluded from all aggregations.
 - a. Foreign Exchange Students (FXS).
 - b. Homeschooled students (schtype='SNE').
 - c. Student in school less than 180 hours (PSNE).
 - d. DNP (for that subject)
 - e. First year in U.S schools LEP*(regardless of how many items were attempted)
 - f. CRT only: Student tested with Non-Standard Accommodations (NSA for that subject)*

g. Alt (Alt='1')

* These students are aggregated on the Disaggregated report in their respective rows.

2. If any of the non-standard accommodations are bubbled the student is considered tested with non-standard accommodations (NSA) in that subject.
3. If the student has not been in that school for the entire academic year the student is excluded from school level aggregations (NSAY).
4. If the student has not been in that system for the entire academic year the student is excluded from system and school level aggregations (NDAY).
5. If the student took the alternate assessment the student is not counted as participating in the general assessment. Alternate Assessment students receive their results on an Alternate Assessment Student Report. They are reported according to participation rules stated in this document.
6. (CRT-Alternate) If the teacher halted the administration of the assessment after the student scored zero (0) for three (3) consecutive items within tasklets , the student is classified as Halted in that subject. If the student was halted within a tasklet then the rest of the items within the tasklet are blanked out and do not count toward the student's score. If the other tasklets are complete then those items will be counted toward the student's score.
7. If the student took the Braille form of the test the raw scores are not included in raw score school, system or state averages. They are not included in group averages on the interactive roster.

D. Student Participation Summary

Participation Status	Part. Flag	Raw score	Scaled Score	Perf. level	Included on Roster	Included in aggregations		
						Sch	Sys	Sta
FXS	E	✓	✓	✓				
SNE	E	✓	✓	✓				
PSNE	E	✓	✓	✓				
NSA(by subject) Applies to CRT only	A	✓	✓	✓	✓	Only included in count and percents on Disaggregated report for nonstandard accommodations.		
First year in U.S schools LEP	A	✓	See Report Specific Rules	See Report Specific Rules	✓			
NSAY only	B	✓	✓	✓	✓		✓	✓

Participation Status	Part. Flag	Raw score	Scaled Score	Perf. level	Included on Roster	Included in aggregations		
						Sch	Sys	Sta
NDAY	C	✓	✓	✓	✓			✓
ALT*	A	✓	✓	✓	✓	See footnote below		
Incomplete	A	✓	✓	✓	✓			
DNP (Non-Participants)	F	✓	✓	✓	✓			
Halted(CRT-Alt only by subject)	D	✓	✓	✓	✓	✓	✓	✓
Tested	Z	✓	✓	✓	✓	✓	✓	✓

* They are included in summary data only for alternate assessment reports (according to participation rules).

If a student has conflicting participation statuses the following hierarchy is applied to determine how the student is reported:

- F (Student attempted no items and is not alt and cannot be classified as first-year LEP)
- E (FXS, SNE or PSNE)
- A (NSA, first year in U.S schools LEP, ALT or INC)
- C (NDAY)
- B (NSAY)
- D (Halted; applies to CRT-Alt only)
- Z (completed CRT or CRT-Alt and none of the above conditions apply)

III. Calculations

A. Raw Scores

1. (CRT) Raw scores are calculated using the scores on common multiple choice and open response items.
2. (CRT-Alternate) Raw score is the sum of the individual item scores.

B. Scaling

1. Scaling is accomplished by defining the unique set of test forms for each grade/subject combination. This is accomplished as follows:
 - a. Translate each form and position into the unique item number assigned to the form/position.
 - b. Order the items by
 - I. Type- multiple choice, short-answer, constructed-response
 - II. Form-common, then by ascending form number.
 - III. Position
 - c. If an item number is on a form, then set the value for that item number to '1', otherwise set to '.'. Set the exception field to '0' to indicate this is an original test form.

- d. If an item number contains an 'X' (item is not included in scaling) then set the item number to '.'. Set the exception field to '1' to indicate this is not an original test form.
 - e. Compress all of the item numbers together into one field in the order defined in step II to create the test for the student.
 - f. Select the distinct set of tests from the student data and order by the exception field and the descending test field.
 - g. Check to see if the test has already been assigned a scale form by looking in the daScaleForm table. If the test exists then assign the existing scale form. Otherwise assign the next available scale form number. All scale form numbering starts at 01 and increments by 1 up to 99.
2. Psychometrics provides a lookup table for each scale form. These lookup tables are used to assign scaled scores, performance levels and standard errors.
 3. The scaled score cuts for all three subjects and all grades have been fixed and are the same as last year for the CRT.
 4. Students excluded from aggregations at the state level are excluded from psychometric files.
- C. CRT-Alternate: The classcode is created using the following steps:
1. The following students are not included when creating the class codes.
 - SNE
 - FXS
 - PSNE
 2. The dataset (by grade) is sorted by schcode and class/group name
 3. The records are then numbered consecutively starting at 1. This number is then padded with zeros (in front) to create a 3 digit number.

D. Performance Level coding:

Numeric Performance Level	Performance level Name	Abbreviation
1(lowest)	Novice	N
2	Nearing Proficient	NP
3	Proficient	P
4(highest)	Advanced	A

E. Rounding Table

Calculation	Rounded (to the nearest)
Static Reports: Percents and averages	Whole number
Item averages : Multiple choice items	The average is multiplied by 100 and rounded to the nearest whole number.
Item averages: Open response items	Open-response item averages are rounded to the nearest tenth.

F. Minimum N size

1. The number of included students (N) in a subject is the number of students in the school/system/state minus FXS minus PRAS minus PRNAS minus PSNE minus SNE minus First year LEP minus Incomplete minus NSA minus DNP.
2. Minimum N size is 10.
3. School/system reports are produced regardless of N-size, except no reports are generated if N=0.

G. The common items are used in reporting the average number of points for each standard.

H. Assignment of rperlevel

1. If the student is marked as taking the CRT-Alt then rperlevel='A', otherwise
2. If the student is classified as did not participate (DNP) then rperlevel='D', otherwise
3. If the student is Incomplete in a subject and not marked first year LEP rperlevel='I', otherwise
4. If the student is incomplete in Reading or has not attempted any items in Reading and is marked first year LEP rperlevel='L' for all subjects, otherwise
5. If the student does not meet any of the above conditions then rperlevel=perlevel.

IV. Report Specific Rules

A. Student Label

1. If a student is First year LEP and incomplete in Reading, the Reading performance level is 'LEP'. The reading scaled score is blank.
2. If a student is First year LEP, the math and science performance levels are the name of the earned performance level and the scaled scores are the student's earned score.
3. If the student is not first year LEP, the performance level name corresponding to the student's earned score is displayed.
4. If the student is First year LEP but is not incomplete in Reading then the student receives his earned scaled score and performance level.
5. If the student is DNP the student receives a student label. The student receives scaled score =200 and performance level=Novice.
6. The student's name is formatted as Lname, Fname.

7. The student's name is uppercase.
8. The school and system names are title case.
9. The labels are sorted alphabetically by Lname, Fname within school and grade.
10. Test date is 2011.
11. Performance level name from section III.D above is shown on the label if the student receives a performance level.

B. Student Report

1. State performance will always appear on the student report, regardless of the student's status.
 - a. A bar on the student report will indicate the percentage of students who appear in each performance level for each subject.
2. If a student is First year LEP and incomplete in Reading, the student will receive the note "Student is Limited English Proficient (LEP). Your student is in his or her first year in a United States school. For further information please contact your school principal or testing director."
3. If the student is First year LEP but is not incomplete in Reading then the student receives his earned scaled score and performance level.
4. If a student is First year LEP, the math and science performance levels are the name of the earned performance level and the scaled score is the student's earned score.
5. If the student is not first year LEP, the performance level name corresponding to the student's earned score is displayed.
6. If the student is incomplete the student receives the scores with the note "Your student did not complete the 2011 CRT. For further information please contact your school principal or testing director."
7. If the student is NSA the student receives his scores with the note "Your student was administered the 2011 CRT with a non-standard testing accommodation. For further information please contact your school principal or testing director."
8. If there is no last name or first name for the student, the name displayed is "Name Not Provided".
9. Alt students who are halted receive their scores and performance level and the note "Teacher halted the administration of one or more of the five tasklets after the student scored a 0 for three consecutive items within a tasklet on two different test administrations. Any completed tasklets have been scored and are reflected in the student's scaled score."
10. If the student is DNP the student receives a Student Report. The student receives scaled score =200 and performance level =Novice. The standards will not be reported. The student receives the note "Student did not participate."
11. If the student had a testing irregularity the student receives the note "A test administration irregularity has affected your student's results. For further information please contact your school principal or testing director."
12. Total Points Possible, Student percent of points earned, and Average state percent are suppressed for students who took Braille test (Braille='1') or who used JAWS (JAWS='1'). This suppression is applied only to the standards which contain the items not on the student's form.

13. For each scored subject, the student report will show a bar with the subject scaled score, as well as an error bar showing the low and high scaled scores, adjusted so these scores are equidistant from the scaled score.
14. Only content standards that apply to the student are printed.
15. The following standards are not reported for either CRT or CRT-Alt:
 - a. Reading standard 3
 - b. Mathematics standard 1
 - c. Science standards 5 and 6

C. Roster & Item Level Report-Alternate Assessment only

1. If a student is First year LEP and the student is not incomplete in Reading:
 - a. The math (and science) performance level is the abbreviation of the earned performance level and the scaled score is the student's earned score.
 - b. The reading performance level is the abbreviation of the earned performance level and the scaled score is the student's earned score.
 - c. The student is excluded from Reading, Math and Science aggregations.
2. If the student is First year LEP and incomplete in Reading
 - a. The student's Reading, Math (and Science) performance levels are 'LEP'
 - b. The student's math (and science) scaled score is the student's earned scaled score and the reading scaled score is blank.
 - c. The student's responses for all subjects are displayed.
 - d. The student is excluded from Math, Reading (and Science) aggregations.
3. If the student is not first year LEP, the performance level abbreviation corresponding to the student's earned score is displayed.
4. If the student is incomplete the student receives the scores with a footnote (†) "Student did not complete the assessment."
5. There is no last name or first name for the student, the name displayed is "Name Not Provided". These students appear at the bottom of the roster.
6. If class/group information is missing the roster is done at the school level.
7. Results for Alternate Assessment students are reported only on their class/group/school's alternate *Roster & Item Level Report*.
8. Within each demonstration school the class is 'DEM'.
9. Only the standards reported on the Summary report are reported on the roster.
10. The student's are sorted by lname, fname
11. Student names are formatted Lname, Fname.
12. Student names are uppercase.
13. Performance level abbreviation from section III.D is placed the performance level column if the student receives are performance level.
14. If the student is NSAY='1' or NDAY='1' then the appropriate footnote is placed beside the first name. ¥ "Not in school and/or system for full academic year."
15. If [subject]halted='1' for any subject then the appropriate footnote is placed beside the first name. § "Teacher halted the administration of one or more of the five tasklets after the student scored a 0 for three consecutive items within

a tasklet on two different test administrations. Any completed tasklets have been scored and are reflected in the student's scaled score."

D. Interactive Roster – CRT only

1. Students who are DNP in a subject are reported with scaled score=200 and performance level='DNP'.
2. Students who are Incomplete in a subject are reported with their earned scaled score and performance level='INC' on the interactive roster.
3. Students who are first-year LEP and who complete the reading test are reported with their earned scaled score and performance level and are included in school, system and state level aggregations for all subjects unless otherwise excluded based on completeness in math or science.
4. Students who are first-year LEP and who do not complete the reading test are reported with their earned scaled score and performance level='LEP' for all subjects. These students are excluded from school, system and state level aggregations.
5. Students who participated in Alternate assessment are listed on the rosters. Their scaled score is blank and the performance level='ALT'. These students are not included in aggregations.
6. The items are reported using the released item number.
7. Students who took the Braille form are not included in any rawscore aggregations. These students have a scaleform other than 01.
8. The following students will have included set to 0 in tblscoreditem (these students are excluded from performance level aggregations):
 - a. The student did not participate in the subject (partstatus='F')
 - b. The student has partstatus='E'
 - c. The student is LEPfirst (LEPfirst='1' regardless of how many items attempted)
 - d. The student is incomplete in the subject.
 - e. The student took the alternate assessment (alt='1')
 - f. Student took the subject with nonstandard accommodations (NSA).
 - g. Student is NSAY (NSAY='1').
 - h. Student is NDAY (NDAY='1').
9. If the student took the Braille form (Braille='1'), included is set to 2. These students are excluded from raw score aggregations.
10. If students do not fall into any of the categories in numbers 8 and 9 above, included is set to '1'.
11. If partstatus='E' for any subject then interactive='0' otherwise interactive='1'. Students with interactive='0' are not available in the interactive site.
12. State level item averages do not include students with school type PRAS, PRNAS or SNE.
13. District level item averages do not include students who are marked nday='1'.
14. Only students whose partstatus is not 'E' for any subject are included in tblStuLongitudinal.
15. The filter column in tblItemAveragesLookup is the concatenation of the gender,ethnic,iep,lep,econdis,migrant and plan504 fields in that order.
16. RepType='0' for all records in tblItemAverages.

E. Summary Report

1. Section I (Distribution of Scores)
 - a. Distribution of Scores will be suppressed and left blank for systems/schools with N less than 10.
2. Section II (Subtest Results) Students with scaleform other than 01 are not included in Subtest Results.
 - a. Subtest Results will be suppressed and left blank for systems/schools with N less than 10.
 - b. A footnote reading “Results are suppressed when less than ten (10) students were assessed.” will appear at the bottom of the first page of the report.
 - c. (Alt only) If any Standard has fewer than 5 Possible Points, the Average Points Earned for that Standard will read “--“ for School, System and State. A footnote will appear below this section reading “-There were too few score points to report on this standard, or no items on the test measured this standard.”
3. Section III (Results for Subgroups of Students)
 - a. Performance level results for subgroups with N less than 10 are suppressed, and the footnote * “Less than 10 students were assessed.” will appear. N is always reported.
 - b. CRT only: Count of students who are considered NSA for that subject excluding those students who are incomplete, nsay (at school level), nday (at school and system level) or FXS or SNE or PSNE or First year LEP or alt (general assessment report).
 - c. Count of First year LEP students excludes those students who are nsay (at school level), nday (at school or system level) or incomplete or FXS or SNE or PSNE or NSA or alt (general assessment).

V. Data File Rules

1. The following students are not included in the state file:
 - a. Alternate Assessment students (in CRT)
 - b. Homeschooled students (SNE)
 - c. Student is in school less than 180 hours (PSNE)
2. If the student receives a performance level ‘LEP’ on the student report in Reading, the student receives LEP for the Reading performance level in the state files.
3. Alt students who are halted are marked ‘1’ in the halted field for that subject.
4. Students who take the Braille form of the test are flagged Braille=’1’ in the state and system level files.
5. In the system level files only the released scored items are included.
6. The following students are not included in the system level files:
 - a. Foreign Exchange students (FXS=’1’)
 - b. Homeschooled students (SNE)
 - c. Student is in school less than 180 hours (PSNE)
7. The following students are not included in the previous year school level files:
 - a. Foreign Exchange students (FXS=’1’)
 - b. Homeschooled students (SNE)
 - c. Student is in school less than 180 hours (PSNE)

VI. PDF file naming conventions to be used by Report Programmer

1. Printed Reports

- a. Labels
MT La [grade].pdf
- b. Student Report (Parent Copy)
#####[systemcode]MT Sr [grade] (Parent Copy).pdf
- c. Student Report (School Copy)
#####[systemcode]MT Sr [grade] (School Copy).pdf

2. Web Reports

- a. School Summary Reports
MT Su Sch [3 character subject][grade].pdf
- b. System Summary Reports
MT Su Dis [3 character subject][grade].pdf
- c. State Summary Reports
MT Su Sta [3 character subject][grade].pdf

VII. Shipping Product Code Summary

1. School (ReportFor='1')

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
03	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
04	Student Labels (CRT)	03	Reading, Math and Science	00	1 set for each school
05	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
06	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
07	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
08	Student Labels (CRT)	03	Reading Math and Science	00	1 set for each school
10	Student Labels (CRT)	03	Reading Math and Science	00	1 set for each school

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
03	Student Report (CRT)	02	Reading and Math	00	1 for each student
04	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
05	Student Report (CRT)	02	Reading Math	00	1 for each student
06	Student Report (CRT)	02	Reading and Math	00	1 for each student
07	Student Report (CRT)	02	Reading and Math	00	1 for each student
08	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
10	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
03	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
04	Student Labels (CRT-Alt)	07	Reading, Math and Science	00	1 set for each school
05	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
06	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
07	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
08	Student Labels (CRT-Alt)	07	Reading Math and Science	00	1 set for each school
10	Student Labels (CRT-Alt)	07	Reading Math and Science	00	1 set for each school
03	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student
04	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student
05	Student Report (CRT-Alt)	08	Reading Math	00	1 for each student
06	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student
07	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student
08	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
10	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student
00	Interp. Guide	04		00	1 per school

Appendix A

1. Items not available on the Braille form

Grade	Form	Content	Positon	Reporting Category
4	01	Reading	14	
4	00	Reading	22	2
4	00	Math	10	4
4	00	Math	43	4
4	01	Math	57	
4	00	Math	72	4
4	00	Science	11	4
4	00	Science	52	2
4	00	Science	55	3
5	01	Reading	67	
5	00	Math	22	4
5	00	Math	25	6
5	01	Math	50	
5	00	Math	72	4
6	01	Math	22	
6	00	Math	73	4

Addenda

Performance Level 2 has been changed from “Nearing Proficient” to “Nearing Proficiency.” The table in Section III.D. of this document should now read as follows:

Performance Level coding:

Numeric Performance Level	Performance level Name	Abbreviation
1(lowest)	Novice	N
2	Nearing Proficiency	NP
3	Proficient	P
4(highest)	Advanced	A

The following rules are in place to clarify suppression rules for CRT-Alt reports:

Summary report:

If the number of students included is less than 10, suppress the aggregate data.

If the number of total possible points is less than 5 for a standard, place a dash in the school, system, and state cells for that standard with a footnote explaining the dash.

Roster Report:

Do not suppress aggregate data regardless of the number of students included.

Do not suppress standard data if the number of total possible points is less than 5.

Student Report:

Do not suppress standard data if the number of total possible points is less than 5.

Aggregate data are at the state level only, so the “less than 10 students included” rule doesn’t apply.

Student Level data files:

Do not suppress standard data if the number of total possible points is less than 5.

Appendix B—PARTICIPATION RATES

Table B-1. 2010–11 Montana CRT: Summary of Participation by Demographic Category – Mathematics

<i>Description</i>	<i>Number tested</i>	<i>Percent</i>
Special Education	7,902	10.84
Title 1	27,874	38.23
Low Income	31,719	43.50
American Indian	8,827	12.11
Asian	702	0.96
Hispanic	2,555	3.50
Black or African American	886	1.22
White, Non-Hispanic	59,716	81.90
Native Hawaiian/Other Pacific Islander	230	0.32
Female	35,406	48.56
Male	37,510	51.44
Limited English Proficient	1,592	2.18
Migrant	150	0.21
Plan 504	556	0.76
All Students	72,916	100.00

Table B-2. 2010–11 Montana CRT: Summary of Participation by Demographic Category – Reading

<i>Description</i>	<i>Number tested</i>	<i>Percent</i>
Special Education	7,889	10.82
Title 1	28,052	38.47
Low Income	31,713	43.49
American Indian	8,818	12.09
Asian	703	0.96
Hispanic	2,556	3.51
Black or African American	885	1.21
White, Non-Hispanic	59,730	81.91
Native Hawaiian/Other Pacific Islander	229	0.31
Female	35,422	48.58
Male	37,499	51.42
Limited English Proficient	1,582	2.17
Migrant	150	0.21
Plan 504	556	0.76
All Students	72,921	100.00

Table B-3. 2010–11 Montana CRT: Summary of Participation by Demographic Category – Science

<i>Description</i>	<i>Number tested</i>	<i>Percent</i>
Special Education	3,339	10.72
Title 1	104	0.33
Low Income	12,719	40.85
American Indian	3,559	11.43
Asian	320	1.03
Hispanic	1,083	3.48
Black or African American	377	1.21
White, Non-Hispanic	25,696	82.53
Native Hawaiian/Other Pacific Islander	100	0.32
Female	15,162	48.70
Male	15,973	51.30
Limited English Proficient	593	1.90
Migrant	55	0.18
Plan 504	300	0.96
All Students	31,135	100.00

Appendix C—ACCOMMODATION FREQUENCIES BY CONTENT AREA

**Table C-1. 2010–11 MontCAS: Numbers of Students Tested With Accommodations by
Accommodation Type and Grade – Mathematics**

<i>Accommodation code</i>	<i>Grade 3</i>	<i>Grade 4</i>	<i>Grade 5</i>	<i>Grade 6</i>	<i>Grade 7</i>	<i>Grade 8</i>	<i>Grade 10</i>
matAccom01	165	207	208	184	100	90	54
matAccom02	263	371	337	311	226	220	126
matAccom04	170	170	180	124	81	84	45
matAccom05	1302	1469	1192	1171	926	915	523
matAccom06	234	257	251	188	80	79	48
matAccom07	808	931	877	654	341	360	176
matAccom08	862	963	820	571	318	367	167
matAccom09	3	2	0	1	2	3	1
matAccom10	2	8	10	9	21	28	3
matAccom12	2	2	0	1	1	0	2
matAccom13	4	5	2	5	1	1	0
matAccom14	1	6	6	0	0	0	0
matAccom15	5	4	3	1	4	2	0
matAccom16	1	5	9	1	0	1	0
matAccom17	0	0	0	0	0	0	0
matAccom18	2	3	1	3	3	3	1
matAccom19	128	169	114	97	37	41	8
matAccom20	6	14	4	8	6	7	0
matAccom21	1	8	0	2	1	1	0
matAccom22	1246	1224	908	744	537	515	231
matAccom23	2	4	9	6	4	6	0
matAccom24	41	57	89	116	36	51	11
matAccom25	105	127	99	123	80	81	23
matAccom26	0	1	2	0	0	0	0
matAccom27	2	5	3	9	7	8	2
matAccom28	0	2	2	3	0	1	0
matAccom30	0	0	0	0	0	0	0
matAccom32	0	0	0	0	0	0	0

Table C-2. 2010–11 MontCAS: Accommodations – Mathematics

<i>Accommodation</i>	<i>Description</i>
matAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
matAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
matAccom04	Individual Administration: Test is administered in a one-to-one situation.
matAccom05	Small Group Administration: Test is administered to a small group of students.
matAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
matAccom07	Alternative Setting: Test is administered to a student in a different setting.
matAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
matAccom09	Home Setting: Test is administered to the student by school personnel in their home.
matAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
matAccom12	Magnification: Student used equipment to magnify test materials.
matAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
matAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
matAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
matAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
matAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
matAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
matAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
matAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
matAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
matAccom22	Oral Presentation: The test administrator must read the test items and answer choices word-for-word. Before reading aloud, the test administrator should advise students that each item and answer choice will be read aloud in exactly the order as presented. Students should also be advised that items, including answer choices, will be repeated at the end of a session in case the students wish to review/check their work.
matAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
matAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
matAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
matAccom26	Braille: Braille version of the test was used by the student.
matAccom27	Large Print: A large-print version of the test is used by student.
matAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
matAccom30	Student uses a calculator, number chart, arithmetic table, or manipulative on the no calculator sections of the mathematics test.
matAccom32	With verification from the OPI in advance of the testing window, some other approved accommodation is used by the student.

**Table C-3. 2010–11 MontCAS: Numbers of Students Tested With Accommodations by
Accommodation Type and Grade – Reading**

<i>Accommodation code</i>	<i>Grade 3</i>	<i>Grade 4</i>	<i>Grade 5</i>	<i>Grade 6</i>	<i>Grade 7</i>	<i>Grade 8</i>	<i>Grade 10</i>
REAAccom01	154	212	204	182	100	91	55
REAAccom02	259	371	357	313	235	244	122
REAAccom04	160	184	190	126	86	87	49
REAAccom05	1277	1412	1166	1152	888	901	550
REAAccom06	221	244	233	176	88	101	50
REAAccom07	785	928	842	648	346	376	186
REAAccom08	851	940	790	566	319	373	170
REAAccom09	3	1	1	1	2	4	1
REAAccom10	2	8	10	9	21	27	3
REAAccom12	2	1	0	1	0	1	2
REAAccom13	8	7	6	5	1	2	0
REAAccom14	4	5	6	1	1	1	0
REAAccom15	4	3	3	1	4	2	0
REAAccom16	1	6	9	7	6	5	5
REAAccom17	0	0	0	0	0	0	0
REAAccom18	2	3	1	3	3	3	1
REAAccom19	161	207	155	133	60	65	13
REAAccom20	6	15	4	10	14	7	6
REAAccom21	1	6	0	3	1	1	1
REAAccom22	1104	1028	794	654	434	430	224
REAAccom23	2	2	10	3	2	2	2
REAAccom24	38	50	88	106	38	59	12
REAAccom25	94	112	91	111	81	83	23
REAAccom26	0	0	2	0	0	0	0
REAAccom27	2	5	3	8	5	7	2
REAAccom28	1	1	5	4	4	1	1
REAAccom29	0	0	0	0	0	0	0
REAAccom31	0	0	0	0	0	0	0

Table C-4. 2010–11 MontCAS: Accommodations – Reading

<i>Accommodation</i>	<i>Description</i>
REAAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
REAAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
REAAccom04	Individual Administration: Test is administered in a one-to-one situation.
REAAccom05	Small Group Administration: Test is administered to a small group of students.
REAAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
REAAccom07	Alternative Setting: Test is administered to a student in a different setting.
REAAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
REAAccom09	Home Setting: Test is administered to the student by school personnel in their home.
REAAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
REAAccom12	Magnification: Student used equipment to magnify test materials.
REAAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
REAAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
REAAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
REAAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
REAAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
REAAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
REAAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
REAAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
REAAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
REAAccom22	Oral Presentation: Only the questions and answer choices may be read aloud to the student. It is advised that the questions be read aloud to the student before he/she reads each passage. After the student has read the passage, the test administrator must read the questions and answer choices word-for-word one at a time in exactly the order as presented.
REAAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
REAAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
REAAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
REAAccom26	Braille: Braille version of the test was used by the student.
REAAccom27	Large Print: A large-print version of the test is used by student.
REAAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
REAAccom29	Reading passages are read aloud to student, or student uses text-reader software for reading passages.
REAAccom31	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.

Table C-5. 2010–11 MontCAS: Numbers of Students Tested With Accommodations by Accommodation Type and Grade – Science

<i>Accommodation code</i>	<i>Grade 4</i>	<i>Grade 8</i>	<i>Grade 10</i>
sciAccom01	214	91	54
sciAccom02	348	216	122
sciAccom04	162	77	54
sciAccom05	1318	916	551
sciAccom06	241	97	46
sciAccom07	900	364	183
sciAccom08	909	357	172
sciAccom09	1	2	1
sciAccom10	9	28	3
sciAccom12	1	1	2
sciAccom13	4	2	0
sciAccom14	6	0	0
sciAccom15	4	2	0
sciAccom16	5	2	1
sciAccom17	0	0	0
sciAccom18	3	3	1
sciAccom19	196	57	13
sciAccom20	14	7	0
sciAccom21	2	1	0
sciAccom22	1103	525	269
sciAccom23	7	6	0
sciAccom24	49	55	12
sciAccom25	132	80	24
sciAccom26	1	0	0
sciAccom27	6	8	2
sciAccom28	2	1	0
sciAccom33	0	0	0

Table C-6. 2010–11 MontCAS: Accommodations – Science

<i>Accommodation</i>	<i>Description</i>
sciAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
sciAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
sciAccom04	Individual Administration: Test is administered in a one-to-one situation.
sciAccom05	Small Group Administration: Test is administered to a small group of students.
sciAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
sciAccom07	Alternative Setting: Test is administered to a student in a different setting.
sciAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
sciAccom09	Home Setting: Test is administered to the student by school personnel in their home.
sciAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
sciAccom12	Magnification: Student used equipment to magnify test materials.
sciAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
sciAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
sciAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
sciAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
sciAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
sciAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
sciAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
sciAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
sciAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
sciAccom22	Oral Presentation: The test administrator must read the test items and answer choices word-for-word and in exactly the order as presented.
sciAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
sciAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
sciAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
sciAccom26	Braille: Braille version of the test was used by the student.
sciAccom27	Large Print: A large-print version of the test is used by student.
sciAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
sciAccom33	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.

Appendix D—ITEM REVIEW COMMITTEE MEMBERS

Table D-1. Passage Review Committee Members—November 4–5, 2009

<i>First name</i>	<i>Last name</i>	<i>Position</i>
Richard	Desch	Curriculum Director
Kris	Goyins	Curriculum Specialist
Keith	Grebetz	Teacher
Dana	Haring	Department Leader for English/Language Arts
Mike	Jetty	Curriculum Specialist
Linda	Jones	Literacy Instructor
Amber	Kleven	English Teacher
Shelly	Moen	4th Grade Teacher
Vicky	Pansuk	Principal/Title I Director
Cory	Pierce	Technology/Dean of Students
Penny	Reynolds	Teacher
Carol	Shipley	History Teacher
Mona	Sindelar	District Testing Coordinator
Patti	Vennes	K-8 Teacher

Table D-2. Benchmarking Committee Members—May 11–13, 2010

<i>First name</i>	<i>Last name</i>	<i>Content</i>
Bette	Paskey	Math
Heidi	Hanks	Math
Beckie	Frisbee	Math
Kevin	Guettler	Math
Michele	Schaub	Reading
Connie	Daily	Reading
Linda	Jones	Reading
Sarah	Hays	Reading
Paul	Tackes	Science
Nina	Miller	Science

Table D-3. Bias Review Committee Members—April 12, 2010

<i>First name</i>	<i>Last name</i>	<i>Position</i>
Zennith	Calfbossribs	Teacher/ Reading Tutor
Richard	Desch	Data/Testing Coordinator
Keith	Grebetz	English Teacher
Jennifer	Gundlach	Special Education Teacher
Mike	Jetty	Indian Education Specialist
Christy	Jobman	Science Teacher
Laura	Monroe	Special Education Teacher
Catherine	Nesmith	Special Education/Title I Teacher
Penny	Reynolds	Teacher
Carol	Shipley	Teacher
Mona	Sindelar	District Testing Coordinator
Chris	Stout	Superintendent
Katrina	Strout	Title I Reading Teacher
Patti	Vennes	Teacher

Table D-4. Item Review Committee Members—April 13–14, 2010

<i>First name</i>	<i>Last name</i>	<i>Position</i>
Michelle	Anderson	English Teacher
Jonna	Ascherman	Third Grade Teacher
Jeffrey	Bennett	4th Grade Teacher
Becky	Berg	Teacher
Lacey	Brunner	English Teacher 8-12
Katie	Burke	Teacher
Susan	Dailey	Resource Teacher 7-12
April	Feeley	Teacher
Kathy	Gaul	4th Grade Teacher
Ashley	Gillespie	8th Grade Physical Science Teacher
Kris	Goyins	Curriculum Specialist
Angela	Haas	HS Science Teacher
Tara	Habel	Teacher
Judy	Haefner	Title I Math Specialist
Heidi	Hanks	6th Grade Teacher
Dana	Haring	Teacher
Linda	Horst	8th Grade Math Teacher
Jean	Howard	Curriculum Specialist
Carol	Idland	Teacher
Alissa	Johnson	Teacher
Shari	Kepner	Math Teacher
Roberta	Kipp	Dean of Students/Language Arts Teacher
Sheryl	Kohl	Math Coach, K-12
Amanda	Kohut	Mathematics
Callie	Kolste	Teacher upper grades
Mike	Lanier	Math Teacher/Head of Math Dept.
Mary	Lyndes	Teacher, Grade 3
Kathleen	Manders	Teacher
Jay	McGraw	Teacher
Ashlee	Mihelish	Teacher
Donald	Miller	English and Business Teacher
Shelly	Moen	4th Grade Teacher
Brian	Moser	High School Math Teacher
Jeremiah	Nitz	Teacher
Jeffery	Padgett	Science Teacher
Janice	Petriz	8th Grade Science Teacher
Kim	Quigley	Teacher
Lynda	Reese	Special Ed. & 6th Grade Math
Michele	Schaub	Teacher
Verne	Schlepp	Retired - 2009
Errin	Schmitz	Curriculum Specialist
Laurie	Schneider	Mathematics Teacher - HS
Constance	Sell	Teacher
Brett	Shelagowski	Teacher
Lynnette	Sheldon	Teacher
Violet	Sinclair	Reading Facilitator
Sue	Solomon	8th Grade Math Teacher
Julie	Stenzel	Teacher
Susan	Sweet	Teacher
Michelle	Swinford	Teacher
Lori	Townsend	Teacher
Susan	Wallace	Teacher
Debra	Westrom	6th Grade Math Teacher
Rodd	Zeiler	Teacher
Angel	Zickefoose	District Math Coach

Table D-5. Item Statistical Review Committee Members—June 29–30, 2010

<i>First name</i>	<i>Last name</i>	<i>Position</i>
Steve	Bell	Science Teacher
Tim	Bolten	Math/Science 8th Grade
Paul	Burckhard	Jr. High/High School Science Teacher
Pam	Diamond	8th Grade Language Arts Teacher
Eleanor	Doucette	Math Teacher, 6th to 12th Grades
Holly	Faris	HS Physical Science Teacher
Kathy	Gaul	4th Grade Teacher
Keith	Grebetz	Middle School English/Reading
Linda	Horst	8th Grade Math Teacher
Carol	Idland	5th Grade Teacher/Librarian
Rolland	Karlin	Teacher 5th Grade
Anne	Keith	8th Grade Algebra/Math Teacher
Sheryl	Kohl	Math Coach, K-12
Callie	Kolste	Supervising Teacher Grades 3–5
Bonnie	Labrie	Title I
Mary	Lyndes	3rd Grade Teacher
Laura	Monroe	Special Education Teacher
Dalene	Normand	3rd Grade Teacher
Jeffery	Padgett	High School Science Teacher
Bette	Paskey	6th Grade Math Teacher
Billie	Perry	HS Science Teacher
Janice	Petriz	Grade 8 Science Teacher
Ellen	Rose	Mathematics 7, 8
Mona	Shortman	Teacher Grade 3
Dana	Haring	Grade 7 English Language Arts
Steve	Bell	Science Teacher
Susan	Solomon	8th Grade Math Teacher
Kris	Goyins	Communication Arts Curriculum Specialist
Margaret	Aukshun	Teacher/Department Chair
Kathje	Dalton	High School Math Teacher
Brian	Kroll	High School Math Teacher
Nina	Miller	Math Specialist
Mike	Lanier	Math Teacher/Head of Math Dept.
Amanda	Kohut	Mathematics Teacher
Michelle	Swinford	High School English Teacher
Becky	Telling	7-8th Language Arts/Reading Teacher
Debbie	Wenger	5th Grade Teacher
Lisa	Wood	Teacher, President Mctm

Appendix E—ITEM LEVEL CLASSICAL STATISTICS

Table E-1. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 3

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
43120	MC	0.94	0.26	0
76764	MC	0.86	0.36	0
35786	MC	0.84	0.42	0
138763	MC	0.82	0.25	0
138780	MC	0.68	0.44	1
138986	MC	0.59	0.47	1
138787	MC	0.69	0.45	1
138879	MC	0.73	0.36	1
76916	MC	0.46	0.47	1
60299	MC	0.73	0.42	1
77013	MC	0.95	0.21	1
138751	MC	0.82	0.39	1
43112	MC	0.89	0.38	1
138865	MC	0.38	0.28	1
76839	MC	0.67	0.22	2
76873	MC	0.85	0.31	2
138929	MC	0.82	0.35	1
60375	MC	0.47	0.36	1
139045	SA	0.86	0.29	0
139051	SA	0.58	0.36	0
138799	CR	0.44	0.54	1
138756	MC	0.94	0.32	0
138917	MC	0.82	0.31	0
59347	MC	0.72	0.44	2
76906	MC	0.69	0.37	0
60271	MC	0.26	0.30	0
138982	MC	0.87	0.44	0
138755	MC	0.69	0.50	1
138765	MC	0.64	0.49	1
43092	MC	0.63	0.38	1
76968	MC	0.44	0.31	1
138832	MC	0.73	0.52	1
42981	MC	0.90	0.33	2
76863	MC	0.53	0.33	2
76920	MC	0.48	0.27	1
76871	MC	0.67	0.30	1
59321	MC	0.70	0.41	1
76780	MC	0.85	0.43	0
138821	MC	0.53	0.32	1
34618	MC	0.79	0.47	1
77041	SA	0.63	0.47	1
138965	MC	0.85	0.26	0
76778	MC	0.83	0.42	1
60276	MC	0.85	0.37	1
76845	MC	0.37	0.41	1
60358	MC	0.63	0.27	0
138781	MC	0.83	0.39	1
138795	MC	0.66	0.34	1
139031	MC	0.56	0.41	1

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
138789	MC	0.49	0.27	3
206058	MC	0.96	0.26	0
139015	MC	0.70	0.40	1
138826	MC	0.27	0.35	1
138958	MC	0.57	0.52	1
138995	MC	0.70	0.45	1
76886	MC	0.79	0.34	1
138824	MC	0.40	0.48	1
139020	MC	0.73	0.52	1
35188	MC	0.93	0.24	1
43277	CR	0.51	0.45	1

Table E-2. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 4

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
76994	MC	0.89	0.31	0
76832	MC	0.79	0.35	0
76789	MC	0.83	0.33	0
77031	MC	0.62	0.46	0
77056	MC	0.84	0.34	0
139961	MC	0.80	0.32	0
139571	MC	0.47	0.24	0
139903	MC	0.26	0.29	0
76822	MC	0.74	0.49	0
139538	MC	0.69	0.38	0
139785	MC	0.56	0.39	0
77024	MC	0.18	0.29	0
62302	MC	0.73	0.27	0
76933	MC	0.55	0.26	0
139763	MC	0.43	0.32	0
34854	MC	0.82	0.31	0
43147	MC	0.81	0.47	0
76804	MC	0.83	0.33	2
140171	SA	0.67	0.44	1
76773	SA	0.45	0.49	1
140183	CR	0.49	0.59	1
76847	MC	0.85	0.35	0
139618	MC	0.88	0.34	0
76803	MC	0.53	0.34	0
62343	MC	0.69	0.21	0
139934	MC	0.57	0.38	0
62317	MC	0.53	0.16	0
62222	MC	0.63	0.46	0
76891	MC	0.74	0.45	0
77026	MC	0.74	0.33	0
139950	MC	0.39	0.31	0
139562	MC	0.75	0.48	0
139542	MC	0.61	0.31	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
62355	MC	0.83	0.47	0
139592	MC	0.63	0.39	0
34824	MC	0.57	0.31	0
34633	MC	0.79	0.44	0
35192	MC	0.59	0.39	0
77046	MC	0.57	0.34	0
43143	MC	0.79	0.35	2
76755	SA	0.76	0.39	0
139915	MC	0.79	0.32	0
139765	MC	0.91	0.38	0
43266	MC	0.63	0.34	0
62324	MC	0.76	0.41	0
139956	MC	0.61	0.44	0
77022	MC	0.67	0.38	0
35222	MC	0.63	0.35	0
77042	MC	0.48	0.39	0
76861	MC	0.67	0.44	0
35220	MC	0.70	0.49	0
139630	MC	0.76	0.39	0
139545	MC	0.58	0.35	0
139911	MC	0.82	0.33	0
34964	MC	0.83	0.31	0
77035	MC	0.47	0.25	0
34960	MC	0.77	0.48	0
61832	MC	0.47	0.40	1
76852	MC	0.91	0.38	1
205515	CR	0.58	0.51	0

Table E-3. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 5

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
77259	MC	0.87	0.37	0
60845	MC	0.75	0.24	0
59861	MC	0.61	0.48	0
43566	MC	0.68	0.45	0
140800	MC	0.32	0.37	0
43581	MC	0.70	0.25	0
59818	MC	0.60	0.53	0
140878	MC	0.47	0.32	0
77309	MC	0.67	0.36	0
140823	MC	0.82	0.32	0
77171	MC	0.77	0.41	0
77247	MC	0.69	0.38	0
140829	MC	0.74	0.32	0
34357	MC	0.57	0.41	0
77207	MC	0.56	0.41	0
77302	MC	0.74	0.29	0
140693	MC	0.57	0.49	0
43437	MC	0.68	0.32	1

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
34605	SA	0.68	0.33	0
77299	SA	0.73	0.47	0
206043	CR	0.56	0.56	1
43600	MC	0.74	0.40	0
140798	MC	0.55	0.41	0
59863	MC	0.86	0.27	0
77254	MC	0.26	0.24	0
140907	MC	0.63	0.24	0
34597	MC	0.37	0.47	0
60383	MC	0.82	0.42	0
43546	MC	0.62	0.32	0
77165	MC	0.49	0.53	0
43465	MC	0.45	0.21	0
43585	MC	0.65	0.34	0
77205	MC	0.79	0.46	0
77228	MC	0.52	0.38	0
43558	MC	0.78	0.47	0
77328	MC	0.63	0.42	0
34696	MC	0.75	0.33	0
140786	MC	0.84	0.42	0
43427	MC	0.56	0.14	0
140951	MC	0.59	0.34	0
62035	SA	0.70	0.42	1
77281	MC	0.82	0.34	0
59995	MC	0.73	0.21	0
34378	MC	0.61	0.48	0
59856	MC	0.61	0.39	0
34367	MC	0.36	0.12	0
34713	MC	0.60	0.28	0
77388	MC	0.74	0.39	0
77169	MC	0.80	0.46	0
77306	MC	0.60	0.49	0
34636	MC	0.34	0.31	0
140919	MC	0.46	0.44	0
43471	MC	0.64	0.34	0
77198	MC	0.45	0.56	0
60398	MC	0.69	0.45	0
34525	MC	0.44	0.48	0
77183	MC	0.79	0.47	0
77241	MC	0.69	0.45	0
77286	MC	0.79	0.28	1
206104	CR	0.47	0.56	1

Table E-4. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 6

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
77340	MC	0.62	0.58	0
77377	MC	0.93	0.27	0
140983	MC	0.63	0.27	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
141156	MC	0.29	0.25	0
77376	MC	0.61	0.46	0
60883	MC	0.55	0.54	0
60890	MC	0.64	0.54	0
43993	MC	0.52	0.35	0
77337	MC	0.37	0.38	0
141165	MC	0.46	0.41	0
43852	MC	0.88	0.42	0
43874	MC	0.52	0.44	0
77317	MC	0.48	0.25	0
43857	MC	0.49	0.47	1
141512	SA	0.72	0.41	0
77646	SA	0.75	0.30	0
43913	SA	0.59	0.51	0
63054	CR	0.38	0.55	3
61130	MC	0.84	0.27	0
43992	MC	0.57	0.17	0
62994	MC	0.45	0.35	0
62073	MC	0.56	0.35	0
141178	MC	0.59	0.41	0
44040	MC	0.66	0.26	0
61136	MC	0.47	0.35	0
77335	MC	0.49	0.42	0
141406	MC	0.73	0.13	0
62060	MC	0.85	0.43	0
61162	MC	0.61	0.36	0
141175	MC	0.57	0.35	0
44027	MC	0.33	0.36	0
62046	MC	0.70	0.30	0
77577	MC	0.68	0.46	0
61151	MC	0.56	0.39	0
77540	MC	0.55	0.23	0
141476	MC	0.71	0.35	0
77414	MC	0.69	0.30	0
141292	MC	0.50	0.41	0
141272	MC	0.38	0.06	0
77621	MC	0.74	0.50	0
77531	MC	0.75	0.47	0
141405	MC	0.51	0.28	0
77461	MC	0.52	0.33	0
141452	MC	0.60	0.31	0
77339	MC	0.49	0.23	0
141283	MC	0.77	0.41	0
34921	MC	0.48	0.38	0
43985	MC	0.67	0.39	0
62047	MC	0.45	0.16	0
43944	MC	0.67	0.44	0
77555	MC	0.74	0.28	0
62044	MC	0.54	0.34	0
141434	MC	0.61	0.43	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
141470	MC	0.66	0.34	0
141407	MC	0.73	0.50	0
140815	MC	0.39	0.26	0
62062	MC	0.69	0.49	0
59842	MC	0.51	0.18	0
62054	MC	0.64	0.35	1
206762	CR	0.26	0.59	1

Table E-5. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 7

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
86269	MC	0.72	0.51	0
142365	MC	0.71	0.46	0
86290	MC	0.80	0.26	0
43791	MC	0.48	0.39	0
86288	MC	0.61	0.35	0
77389	MC	0.41	0.38	0
142368	MC	0.65	0.36	0
86264	MC	0.48	0.46	0
86330	MC	0.57	0.48	0
86273	MC	0.57	0.30	0
86302	MC	0.52	0.44	0
43836	MC	0.59	0.36	0
43862	MC	0.43	0.37	0
43842	MC	0.91	0.35	1
43911	SA	0.52	0.51	1
61374	SA	0.69	0.48	1
86339	SA	0.54	0.54	1
142673	CR	0.34	0.61	3
61766	MC	0.88	0.30	0
86624	MC	0.69	0.33	0
142680	MC	0.39	0.30	0
61244	MC	0.84	0.32	0
86382	MC	0.79	0.40	0
142645	MC	0.70	0.46	0
142735	MC	0.28	0.26	0
142377	MC	0.39	0.47	0
86549	MC	0.55	0.43	0
86606	MC	0.56	0.46	0
86687	MC	0.47	0.42	0
86672	MC	0.72	0.39	0
142376	MC	0.56	0.48	0
61781	MC	0.74	0.43	0
61363	MC	0.56	0.29	0
86597	MC	0.56	0.38	0
43654	MC	0.65	0.36	0
86568	MC	0.52	0.35	0
61279	MC	0.87	0.44	0
43753	MC	0.57	0.50	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
86431	MC	0.42	0.30	0
86535	MC	0.79	0.29	0
86700	MC	0.32	0.25	0
61799	MC	0.66	0.30	0
61874	MC	0.65	0.41	0
86681	MC	0.85	0.41	0
86445	MC	0.38	0.38	0
43680	MC	0.29	0.27	0
86333	MC	0.59	0.38	0
43767	MC	0.84	0.23	0
142818	MC	0.37	0.20	0
61745	MC	0.65	0.44	0
142649	MC	0.65	0.48	0
44173	MC	0.23	0.22	0
61348	MC	0.54	0.45	0
86644	MC	0.57	0.23	0
61772	MC	0.90	0.40	0
86574	MC	0.62	0.35	0
86641	MC	0.63	0.46	0
43780	MC	0.72	0.45	0
43719	MC	0.60	0.45	1
43828	CR	0.56	0.49	1

Table E-6. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 8

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
144251	MC	0.69	0.40	0
144191	MC	0.68	0.28	0
87653	MC	0.49	0.41	0
87576	MC	0.45	0.25	0
87566	MC	0.35	0.27	0
88026	MC	0.79	0.36	0
87571	MC	0.58	0.41	0
43840	MC	0.38	0.42	0
35025	MC	0.72	0.45	0
44181	MC	0.53	0.39	0
63242	MC	0.62	0.43	0
61198	MC	0.38	0.36	0
34979	MC	0.42	0.35	1
44209	MC	0.62	0.43	1
34939	SA	0.58	0.44	1
34935	SA	0.51	0.39	4
34937	SA	0.43	0.51	1
87874	CR	0.50	0.64	5
88177	MC	0.76	0.31	0
44127	MC	0.82	0.34	0
34978	MC	0.59	0.26	0
35062	MC	0.50	0.37	0
35019	MC	0.90	0.38	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
44213	MC	0.31	0.29	0
43888	MC	0.59	0.37	0
144456	MC	0.73	0.36	0
144499	MC	0.62	0.23	0
63048	MC	0.73	0.47	0
63239	MC	0.56	0.35	0
88892	MC	0.53	0.31	0
88277	MC	0.43	0.22	0
63109	MC	0.40	0.21	0
88838	MC	0.63	0.31	0
44633	MC	0.56	0.42	0
144905	MC	0.77	0.31	0
34969	MC	0.29	0.25	0
34932	MC	0.37	0.27	0
88352	MC	0.70	0.36	0
88216	MC	0.72	0.40	0
34995	MC	0.95	0.28	0
63038	MC	0.81	0.45	0
144955	MC	0.70	0.32	0
144942	MC	0.54	0.28	0
144436	MC	0.47	0.36	0
34982	MC	0.64	0.39	0
88361	MC	0.53	0.36	0
63115	MC	0.46	0.28	0
44210	MC	0.64	0.38	0
88313	MC	0.51	0.29	0
88086	MC	0.52	0.38	0
63066	MC	0.39	0.37	0
86542	MC	0.67	0.30	0
44175	MC	0.31	0.32	0
44154	MC	0.49	0.30	0
88187	MC	0.41	0.36	0
88331	MC	0.51	0.54	0
34933	MC	0.55	0.46	0
63256	MC	0.70	0.39	0
63106	MC	0.82	0.31	0
205490	CR	0.46	0.62	1

Table E-7. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 10

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
146070	MC	0.57	0.38	0
77592	MC	0.44	0.45	0
77369	MC	0.64	0.18	0
144377	MC	0.63	0.29	0
43725	MC	0.26	0.28	0
77512	MC	0.34	0.46	0
77364	MC	0.85	0.39	0
59403	MC	0.65	0.29	0
144570	MC	0.46	0.53	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
77501	MC	0.39	0.33	0
145022	MC	0.50	0.26	0
59371	MC	0.30	0.30	0
34804	MC	0.53	0.30	0
43687	MC	0.41	0.29	0
77644	SA	0.25	0.53	5
145015	SA	0.68	0.46	2
77629	SA	0.41	0.54	2
62391	CR	0.25	0.63	11
62361	MC	0.61	0.51	0
62279	MC	0.70	0.36	0
144996	MC	0.49	0.41	0
145032	MC	0.52	0.35	0
146556	MC	0.42	0.22	0
146573	MC	0.39	0.31	0
43951	MC	0.32	0.26	0
77546	MC	0.43	0.58	0
145719	MC	0.53	0.27	0
144825	MC	0.56	0.47	0
145637	MC	0.58	0.23	0
77428	MC	0.63	0.42	0
77374	MC	0.27	0.40	0
77602	MC	0.16	0.10	0
146535	MC	0.49	0.38	0
77358	MC	0.58	0.40	0
146539	MC	0.51	0.47	0
77420	MC	0.39	0.34	0
77424	MC	0.29	0.30	0
144896	MC	0.55	0.40	0
62380	MC	0.53	0.17	0
145292	MC	0.57	0.41	0
43969	MC	0.87	0.28	0
61301	MC	0.50	0.16	0
77430	MC	0.49	0.46	0
144841	MC	0.54	0.40	0
43717	MC	0.73	0.39	0
77544	MC	0.61	0.55	0
34761	MC	0.61	0.48	0
146572	MC	0.51	0.53	0
43631	MC	0.25	0.21	0
146580	MC	0.56	0.46	0
35229	MC	0.38	0.32	0
61324	MC	0.87	0.35	0
62319	MC	0.48	0.47	0
61305	MC	0.30	0.21	1
62365	MC	0.41	0.12	1
144904	MC	0.70	0.41	0
77435	MC	0.33	0.14	0
77503	MC	0.41	0.10	1
77432	MC	0.63	0.22	1
59407	CR	0.35	0.70	3

Table E-8. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 3

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
150711	MC	0.66	0.23	0
150713	MC	0.56	0.26	1
150718	MC	0.85	0.38	0
150721	MC	0.88	0.39	1
150725	MC	0.89	0.38	2
150869	MC	0.93	0.33	0
150951	MC	0.89	0.33	0
151173	MC	0.69	0.46	0
151176	MC	0.87	0.46	0
151186	MC	0.76	0.31	0
151193	MC	0.55	0.42	1
151194	MC	0.50	0.36	1
151200	MC	0.80	0.42	2
151203	MC	0.77	0.41	0
151207	MC	0.48	0.33	0
151212	MC	0.65	0.42	1
151215	MC	0.81	0.41	1
151227	MC	0.57	0.34	1
153156	MC	0.57	0.38	1
151240	CR	0.35	0.55	2
67148	MC	0.74	0.41	0
67151	MC	0.78	0.49	0
67155	MC	0.87	0.36	0
67167	MC	0.78	0.34	1
67184	MC	0.65	0.30	0
67193	MC	0.73	0.52	1
67198	MC	0.49	0.32	4
151127	MC	0.92	0.31	0
151131	MC	0.60	0.29	2
151133	MC	0.70	0.38	0
151141	MC	0.58	0.40	1
151135	MC	0.75	0.49	1
151156	MC	0.83	0.48	1
151166	MC	0.48	0.27	1
150634	MC	0.46	0.36	0
150644	MC	0.59	0.35	0
150648	MC	0.83	0.31	1
150653	MC	0.88	0.33	1
150656	MC	0.60	0.32	2
150664	MC	0.86	0.39	0
150670	MC	0.79	0.49	1
150953	MC	0.77	0.40	1
150962	MC	0.81	0.55	1
150969	MC	0.75	0.44	1
150971	MC	0.90	0.43	1
150974	MC	0.73	0.51	1
150975	MC	0.91	0.45	1
150977	MC	0.71	0.49	1
150978	MC	0.90	0.47	1

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
150982	MC	0.50	0.31	1
150983	MC	0.80	0.42	2
150986	MC	0.67	0.34	0
150988	MC	0.48	0.38	1
150991	CR	0.41	0.49	1

Table E-9. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 4

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
67140	MC	0.73	0.36	0
67149	MC	0.88	0.26	0
67154	MC	0.85	0.33	0
67158	MC	0.59	0.41	0
67164	MC	0.77	0.48	0
67160	MC	0.79	0.34	0
67166	MC	0.80	0.20	0
151672	MC	0.87	0.35	0
151679	MC	0.56	0.32	0
151683	MC	0.78	0.46	0
151684	MC	0.63	0.32	0
151686	MC	0.69	0.35	0
151695	MC	0.79	0.41	0
151697	MC	0.83	0.45	0
151740	MC	0.80	0.37	0
151743	MC	0.88	0.34	0
151745	MC	0.59	0.35	1
151751	MC	0.75	0.50	1
151762	MC	0.58	0.37	1
151772	CR	0.45	0.43	1
151780	MC	0.40	0.38	0
151782	MC	0.88	0.37	0
151783	MC	0.86	0.30	0
151788	MC	0.68	0.34	0
151790	MC	0.69	0.35	0
151794	MC	0.63	0.34	0
151800	MC	0.76	0.45	0
66546	MC	0.51	0.21	0
66550	MC	0.84	0.42	0
66555	MC	0.59	0.39	0
66568	MC	0.79	0.36	0
66571	MC	0.71	0.35	0
66578	MC	0.48	0.48	0
66586	MC	0.70	0.33	0
66597	MC	0.85	0.38	0
66612	MC	0.69	0.30	0
66650	MC	0.84	0.40	0
66615	MC	0.87	0.45	0
66629	MC	0.78	0.38	0
66627	MC	0.84	0.43	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
66634	MC	0.70	0.45	0
151597	MC	0.70	0.42	0
151612	MC	0.83	0.31	0
151615	MC	0.84	0.53	0
151621	MC	0.60	0.36	0
151626	MC	0.84	0.48	0
151632	MC	0.52	0.38	0
151635	MC	0.71	0.44	0
151637	MC	0.83	0.38	0
151655	MC	0.60	0.33	0
151638	MC	0.73	0.41	1
151639	MC	0.62	0.38	0
151644	MC	0.68	0.42	1
151659	CR	0.34	0.45	1

Table E-10. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 5

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
150580	MC	0.89	0.38	0
150578	MC	0.69	0.42	0
150586	MC	0.78	0.18	0
150591	MC	0.86	0.38	0
150593	MC	0.68	0.42	0
150596	MC	0.58	0.34	0
150594	MC	0.79	0.42	0
150415	MC	0.85	0.28	0
150413	MC	0.66	0.40	0
150417	MC	0.83	0.41	0
150419	MC	0.74	0.44	0
200624	MC	0.78	0.36	0
150425	MC	0.81	0.30	0
150427	MC	0.89	0.42	0
150423	MC	0.73	0.32	0
150432	MC	0.51	0.33	0
150440	MC	0.90	0.41	1
150442	MC	0.77	0.30	0
150450	MC	0.42	0.24	1
150456	CR	0.47	0.45	1
65580	MC	0.54	0.37	0
65593	MC	0.78	0.33	0
65616	MC	0.84	0.33	0
65599	MC	0.65	0.45	0
65607	MC	0.69	0.14	0
65611	MC	0.63	0.36	0
65613	MC	0.60	0.36	0
155431	MC	0.74	0.33	0
150527	MC	0.62	0.39	0
150536	MC	0.65	0.43	0
150530	MC	0.88	0.42	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
150547	MC	0.80	0.41	0
150548	MC	0.56	0.32	0
150551	MC	0.68	0.35	0
41321	MC	0.70	0.16	0
41322	MC	0.47	0.37	0
41323	MC	0.74	0.36	0
41326	MC	0.81	0.47	0
41328	MC	0.69	0.45	0
41329	MC	0.63	0.34	0
41331	MC	0.82	0.43	0
150470	MC	0.90	0.41	0
150480	MC	0.88	0.47	0
150474	MC	0.77	0.48	0
150483	MC	0.82	0.40	0
150485	MC	0.74	0.43	0
150488	MC	0.86	0.37	0
150489	MC	0.72	0.48	0
150491	MC	0.79	0.46	0
150492	MC	0.81	0.39	0
150493	MC	0.80	0.43	0
150494	MC	0.80	0.45	0
150505	MC	0.67	0.36	1
150518	CR	0.50	0.39	0

Table E-11. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 6

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
42118	MC	0.90	0.29	0
42131	MC	0.82	0.42	0
42130	MC	0.92	0.42	0
42282	MC	0.66	0.48	0
42132	MC	0.77	0.36	0
42287	MC	0.90	0.37	0
42284	MC	0.71	0.38	0
95414	MC	0.85	0.38	0
95417	MC	0.75	0.27	0
95422	MC	0.86	0.38	0
95426	MC	0.48	0.26	0
95447	MC	0.87	0.40	0
95429	MC	0.88	0.47	0
95446	MC	0.71	0.51	0
95456	MC	0.46	0.25	0
95448	MC	0.80	0.41	0
95451	MC	0.65	0.37	0
95459	MC	0.69	0.43	0
95455	MC	0.78	0.41	0
95469	CR	0.45	0.48	1
151296	MC	0.88	0.41	0
151295	MC	0.71	0.35	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
151278	MC	0.75	0.39	0
151274	MC	0.65	0.35	0
151273	MC	0.85	0.31	0
151282	MC	0.85	0.28	0
151302	MC	0.52	0.34	0
151312	MC	0.62	0.36	0
151315	MC	0.72	0.36	0
151321	MC	0.60	0.23	0
151325	MC	0.60	0.36	0
151337	MC	0.79	0.38	0
151338	MC	0.52	0.35	0
151341	MC	0.73	0.43	0
42100	MC	0.91	0.35	0
42099	MC	0.81	0.41	0
42102	MC	0.76	0.42	0
201077	MC	0.87	0.41	0
42104	MC	0.86	0.43	0
201078	MC	0.72	0.41	0
42106	MC	0.61	0.36	0
68240	MC	0.53	0.35	0
68242	MC	0.63	0.29	0
68247	MC	0.69	0.37	0
68250	MC	0.45	0.35	0
68260	MC	0.70	0.49	0
68263	MC	0.66	0.50	0
68291	MC	0.46	0.32	0
68283	MC	0.75	0.41	0
201084	MC	0.55	0.34	0
68265	MC	0.71	0.46	1
68285	MC	0.68	0.24	0
68269	MC	0.47	0.27	0
68294	CR	0.41	0.48	0

Table E-12. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 7

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
149061	MC	0.68	0.36	0
149064	MC	0.79	0.38	0
149066	MC	0.86	0.51	0
154487	MC	0.86	0.35	0
200972	MC	0.87	0.43	0
200973	MC	0.69	0.48	0
149080	MC	0.77	0.45	0
148792	MC	0.76	0.33	0
148807	MC	0.69	0.46	0
148811	MC	0.67	0.50	0
148827	MC	0.66	0.36	0
148834	MC	0.82	0.48	0
148839	MC	0.90	0.48	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
148843	MC	0.71	0.41	0
148849	MC	0.73	0.31	0
148854	MC	0.83	0.53	0
200974	MC	0.86	0.49	0
148875	MC	0.81	0.50	0
148883	MC	0.79	0.41	0
148887	CR	0.53	0.55	1
92453	MC	0.56	0.46	0
92455	MC	0.86	0.46	0
92462	MC	0.63	0.46	0
92464	MC	0.78	0.37	0
92458	MC	0.71	0.47	0
92471	MC	0.80	0.42	0
92472	MC	0.75	0.34	0
92359	MC	0.80	0.29	0
92384	MC	0.91	0.35	0
92399	MC	0.75	0.46	0
92391	MC	0.84	0.44	0
92395	MC	0.69	0.47	0
92397	MC	0.85	0.43	0
92404	MC	0.80	0.51	0
149083	MC	0.76	0.37	0
149088	MC	0.72	0.40	0
149089	MC	0.80	0.40	0
149090	MC	0.76	0.38	0
149091	MC	0.80	0.36	0
149093	MC	0.61	0.35	0
149098	MC	0.59	0.36	0
148965	MC	0.76	0.41	0
200975	MC	0.68	0.36	0
148970	MC	0.78	0.43	0
148972	MC	0.82	0.46	0
148973	MC	0.84	0.50	0
148977	MC	0.65	0.34	0
148982	MC	0.58	0.39	0
148988	MC	0.78	0.44	0
148993	MC	0.48	0.34	0
149003	MC	0.72	0.37	1
149007	MC	0.83	0.50	0
149008	MC	0.68	0.32	0
149016	CR	0.49	0.61	1

Table E-13. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 8

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
149129	MC	0.68	0.48	0
149130	MC	0.55	0.47	0
149134	MC	0.73	0.24	0
149136	MC	0.63	0.44	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
149139	MC	0.76	0.37	0
152410	MC	0.76	0.43	0
149144	MC	0.81	0.40	0
149342	MC	0.60	0.36	0
149344	MC	0.76	0.44	0
149347	MC	0.65	0.50	0
152841	MC	0.77	0.45	0
149349	MC	0.61	0.43	0
149352	MC	0.79	0.41	0
149353	MC	0.75	0.34	0
149354	MC	0.62	0.39	0
149355	MC	0.61	0.46	0
149360	MC	0.80	0.42	0
149359	MC	0.78	0.46	0
149362	MC	0.91	0.37	0
149368	CR	0.58	0.60	0
149326	MC	0.67	0.32	0
149328	MC	0.84	0.36	0
149331	MC	0.75	0.34	0
149339	MC	0.80	0.46	0
149340	MC	0.75	0.45	0
149341	MC	0.66	0.38	0
149335	MC	0.90	0.37	0
153228	MC	0.84	0.39	0
149261	MC	0.71	0.26	0
149262	MC	0.62	0.27	0
149269	MC	0.85	0.44	0
149276	MC	0.69	0.47	0
149281	MC	0.61	0.45	0
149287	MC	0.69	0.46	0
149372	MC	0.55	0.33	0
149388	MC	0.68	0.47	0
153158	MC	0.68	0.28	0
149377	MC	0.65	0.40	0
149379	MC	0.82	0.50	0
149380	MC	0.80	0.46	0
149383	MC	0.75	0.50	0
149174	MC	0.62	0.40	0
149188	MC	0.76	0.48	0
149165	MC	0.60	0.39	0
149170	MC	0.84	0.54	0
149171	MC	0.86	0.43	0
156089	MC	0.62	0.37	0
149176	MC	0.73	0.39	0
149179	MC	0.85	0.50	0
149181	MC	0.79	0.47	0
149187	MC	0.77	0.41	0
149190	MC	0.73	0.42	0
149191	MC	0.77	0.39	0
149193	CR	0.59	0.58	1

Table E-14. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 10

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
149467	MC	0.87	0.31	0
149468	MC	0.68	0.39	0
149471	MC	0.73	0.29	0
149472	MC	0.76	0.37	0
149482	MC	0.85	0.33	0
149483	MC	0.79	0.39	0
149486	MC	0.85	0.33	0
149545	MC	0.72	0.41	0
149548	MC	0.83	0.39	0
149558	MC	0.71	0.24	0
149544	MC	0.85	0.47	0
149551	MC	0.56	0.31	0
149550	MC	0.73	0.42	0
149549	MC	0.70	0.30	0
149560	MC	0.66	0.51	0
149555	MC	0.83	0.47	0
149554	MC	0.76	0.44	0
149556	MC	0.81	0.48	0
149564	MC	0.76	0.47	0
149566	CR	0.51	0.52	1
42224	MC	0.67	0.35	0
42225	MC	0.73	0.33	0
42226	MC	0.79	0.34	0
42228	MC	0.74	0.44	0
42230	MC	0.65	0.39	0
42233	MC	0.80	0.40	0
42232	MC	0.72	0.34	0
149894	MC	0.74	0.34	0
149896	MC	0.78	0.40	0
149869	MC	0.62	0.29	0
149883	MC	0.86	0.44	0
149887	MC	0.59	0.24	0
149905	MC	0.80	0.32	0
149950	MC	0.90	0.38	0
149515	MC	0.69	0.38	0
149521	MC	0.76	0.36	0
149525	MC	0.75	0.45	0
149528	MC	0.74	0.35	0
149538	MC	0.79	0.43	0
149537	MC	0.83	0.45	0
149541	MC	0.85	0.43	0
149610	MC	0.57	0.32	0
149611	MC	0.84	0.47	0
149613	MC	0.46	0.40	0
149615	MC	0.69	0.39	0
149616	MC	0.65	0.49	0
149627	MC	0.82	0.50	0
149633	MC	0.63	0.40	0
149623	MC	0.50	0.39	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
149629	MC	0.76	0.47	0
149638	MC	0.64	0.34	1
149630	MC	0.46	0.35	0
149631	MC	0.70	0.30	0
149648	CR	0.47	0.62	3

Table E-15. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 4

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
39203	MC	0.57	0.32	0
120003	MC	0.66	0.40	0
75718	MC	0.68	0.27	0
75719	MC	0.58	0.40	0
42780	MC	0.88	0.20	0
120024	MC	0.50	0.40	0
38537	MC	0.74	0.28	0
119986	MC	0.70	0.30	0
75818	MC	0.95	0.24	0
120106	MC	0.39	0.24	0
76397	MC	0.41	0.23	0
142358	MC	0.52	0.40	0
134838	MC	0.65	0.25	0
39164	MC	0.78	0.34	0
75708	MC	0.71	0.29	0
76401	MC	0.61	0.29	0
75831	MC	0.84	0.40	1
75757	CR	0.61	0.49	1
134736	MC	0.76	0.43	0
134742	MC	0.70	0.14	0
120564	MC	0.96	0.24	0
120617	MC	0.83	0.34	0
119971	MC	0.87	0.32	0
134858	MC	0.82	0.36	0
52579	MC	0.92	0.27	0
119973	MC	0.60	0.38	0
57902	MC	0.72	0.37	0
75760	MC	0.79	0.31	0
75691	MC	0.87	0.34	0
76289	MC	0.81	0.40	0
75692	MC	0.57	0.33	0
56263	MC	0.49	0.18	0
144837	MC	0.51	0.30	0
53280	MC	0.82	0.26	0
75693	MC	0.80	0.45	0
57920	MC	0.48	0.17	1
75774	MC	0.51	0.39	0
75695	MC	0.89	0.23	0
120543	MC	0.58	0.29	0
120162	MC	0.63	0.40	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
134745	MC	0.49	0.30	0
75424	MC	0.69	0.31	0
120014	MC	0.73	0.29	0
75406	MC	0.58	0.25	0
75513	MC	0.72	0.31	0
134675	MC	0.68	0.42	0
55442	MC	0.55	0.41	0
56367	MC	0.61	0.31	0
75405	MC	0.62	0.33	0
56225	MC	0.50	0.34	0
56327	MC	0.56	0.36	0
75412	MC	0.69	0.47	0
75404	MC	0.61	0.14	0
75399	MC	0.38	0.29	1
120312	CR	0.47	0.41	1

Table E-16. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 8

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
89277	MC	0.93	0.36	0
89752	MC	0.72	0.30	0
89591	MC	0.57	0.33	0
54228	MC	0.90	0.25	0
89859	MC	0.40	0.17	0
122023	MC	0.67	0.38	0
121232	MC	0.61	0.32	0
89392	MC	0.37	0.21	0
89361	MC	0.66	0.37	0
39587	MC	0.70	0.19	0
89900	MC	0.61	0.38	0
56879	MC	0.44	0.10	0
144065	MC	0.74	0.41	0
56773	MC	0.79	0.34	0
89504	MC	0.71	0.39	0
89892	MC	0.67	0.40	0
56876	MC	0.58	0.45	0
89534	CR	0.23	0.60	1
89433	MC	0.56	0.18	0
56989	MC	0.36	0.14	0
89644	MC	0.52	0.32	0
39701	MC	0.62	0.31	0
122738	MC	0.49	0.26	0
89860	MC	0.29	0.17	0
56846	MC	0.68	0.28	0
89637	MC	0.50	0.25	0
89917	MC	0.84	0.45	0
89934	MC	0.54	0.26	0
39729	MC	0.46	0.30	0
39773	MC	0.52	0.34	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
89895	MC	0.83	0.39	0
122736	MC	0.69	0.48	0
89702	MC	0.72	0.37	0
39648	MC	0.85	0.39	0
89426	MC	0.64	0.38	0
89901	MC	0.49	0.23	0
89857	MC	0.65	0.28	0
121209	MC	0.44	0.20	0
89621	MC	0.46	0.26	0
89406	MC	0.71	0.47	0
56985	MC	0.50	0.35	0
54237	MC	0.74	0.39	0
89463	MC	0.34	0.31	0
39782	MC	0.59	0.49	0
56894	MC	0.80	0.37	0
89710	MC	0.38	0.17	0
89881	MC	0.58	0.32	0
121222	MC	0.68	0.42	0
122036	MC	0.83	0.35	0
39780	MC	0.88	0.40	0
89914	MC	0.42	0.25	0
89513	MC	0.83	0.27	0
39546	MC	0.28	0.21	0
89766	MC	0.62	0.35	0
89958	CR	0.39	0.55	2

Table E-17. 2010–11 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 10

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
134543	MC	0.60	0.26	0
134489	MC	0.73	0.41	0
197896	MC	0.61	0.27	0
119676	MC	0.40	0.31	0
75644	MC	0.42	0.23	0
75433	MC	0.54	0.38	0
75441	MC	0.55	0.30	0
134482	MC	0.77	0.43	0
75865	MC	0.51	0.21	0
120311	MC	0.63	0.40	0
119661	MC	0.59	0.29	0
130550	MC	0.88	0.41	0
134545	MC	0.61	0.21	0
75786	MC	0.73	0.12	0
40188	MC	0.39	0.21	0
134487	MC	0.84	0.42	0
75976	MC	0.78	0.37	0
119960	CR	0.39	0.56	2
75963	MC	0.58	0.28	0
134480	MC	0.51	0.34	0

continued

<i>Item number</i>	<i>Item type</i>	<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent omitted</i>
75875	MC	0.81	0.46	0
119955	MC	0.51	0.35	0
120064	MC	0.78	0.31	0
75859	MC	0.60	0.28	0
134799	MC	0.77	0.45	0
75635	MC	0.53	0.29	0
75941	MC	0.70	0.36	0
75725	MC	0.37	0.23	0
134778	MC	0.38	0.24	0
119674	MC	0.55	0.25	0
75803	MC	0.57	0.34	0
134770	MC	0.61	0.53	0
119664	MC	0.47	0.28	0
75452	MC	0.51	0.33	0
75632	MC	0.51	0.39	0
52960	MC	0.44	0.18	0
119943	MC	0.69	0.37	0
75787	MC	0.80	0.37	0
134792	MC	0.72	0.39	0
75620	MC	0.38	0.32	0
75804	MC	0.49	0.41	0
75726	MC	0.40	0.15	0
134479	MC	0.55	0.40	0
119675	MC	0.43	0.26	0
130594	MC	0.58	0.30	0
75863	MC	0.65	0.47	0
130592	MC	0.64	0.45	0
75634	MC	0.46	0.45	0
75854	MC	0.66	0.29	0
119656	MC	0.60	0.42	0
75630	MC	0.47	0.18	0
75970	MC	0.53	0.41	0
75777	MC	0.77	0.44	0
75846	MC	0.72	0.34	0
52953	CR	0.32	0.46	2

Appendix F—ITEM LEVEL SCORE DISTRIBUTIONS

Table F-1. 2010–11 MontCAS: Item Level Score Distributions for Constructed-Response Items by Subject and Grade

Subject	Grade	Total possible points	Item number	Percent of students at score point				
				0	1	2	3	4
Mathematics	3	4	138799	21.49	27.48	17.53	16.77	15.90
		4	43277	18.47	18.30	20.47	22.11	19.72
	4	4	140183	14.41	21.28	32.58	11.98	18.42
		4	205515	18.08	7.77	11.44	48.61	13.69
	5	4	206043	7.91	18.68	24.34	34.38	13.62
		4	206104	24.81	20.90	14.84	17.70	21.19
	6	4	63054	16.85	37.59	25.69	4.81	11.70
		4	206762	42.14	25.55	18.87	8.80	4.08
	7	4	142673	21.89	31.06	28.64	10.86	4.17
		4	43828	12.32	8.94	26.28	42.79	8.74
	8	4	87874	16.32	9.52	34.35	17.81	17.43
		4	205490	14.24	32.29	25.31	8.22	18.93
	10	4	62391	32.75	29.17	15.15	8.45	3.54
		4	59407	31.20	27.25	19.17	5.23	14.27
Reading	3	4	151240	14.85	43.73	26.05	9.21	4.46
		4	150991	13.23	29.51	37.84	14.89	3.44
	4	4	151772	3.64	28.00	52.01	13.61	1.75
		4	151659	11.39	44.34	37.80	5.25	0.48
	5	4	150456	4.95	30.46	42.09	16.11	5.76
		4	150518	2.79	29.35	40.39	19.39	7.61
	6	4	95469	5.62	29.95	42.87	18.03	2.89
		4	68294	8.89	38.08	35.62	14.26	2.73
	7	4	148887	4.90	20.28	37.52	28.20	8.45
		4	149016	9.04	26.85	30.19	24.18	9.20
	8	4	149368	6.88	17.25	28.29	32.00	15.22
		4	149193	4.91	16.79	30.81	30.25	16.47
	10	4	149566	6.38	17.83	43.67	24.03	6.68
		4	149648	12.53	20.24	34.08	22.30	7.89
Science	4	4	75757	7.59	20.18	22.83	17.10	31.75
		4	120312	11.34	12.86	55.34	15.15	4.78
	8	4	89534	39.75	34.13	17.17	6.42	1.34
		4	89958	23.92	22.16	27.32	16.67	7.59
	10	4	119960	21.39	22.51	34.71	13.24	5.69
		4	52953	12.46	54.41	22.34	6.42	1.99

Appendix G—NUMBER OF ITEMS CLASSIFIED INTO DIF CATEGORIES

**Table G-1. 2010–11 MontCAS: Number of Items Classified as “Low” or “High” DIF,
Overall and by Group Favored – Mathematics**

Grade	Reference group	Focal group	Item type	Number of items	Number “low”			Number “high”		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
3	Male	Female	MC	55	3	3	0	0	0	0
	White	Hispanic	MC	55	5	3	2	0	0	0
	White	Native American	MC	55	2	2	0	2	2	0
	No Disability	Disability	MC	55	2	2	0	0	0	0
	Not Low Income	Low Income	MC	55	2	2	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	11	10	1	6	6	0
	Male	Female	OR	5	0	0	0	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	0	0	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	1	0	1	0	0	0
4	Male	Female	MC	55	5	3	2	0	0	0
	White	Hispanic	MC	55	5	3	2	0	0	0
	White	Native American	MC	55	4	4	0	0	0	0
	No Disability	Disability	MC	55	2	2	0	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	11	7	4	7	6	1
	Male	Female	OR	5	0	0	0	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	0	0	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	2	2	0	0	0	0
5	Male	Female	MC	55	4	2	2	1	1	0
	White	Hispanic	MC	55	2	2	0	0	0	0
	White	Native American	MC	55	2	2	0	0	0	0
	No Disability	Disability	MC	55	7	7	0	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	13	7	6	6	6	0
	Male	Female	OR	5	1	0	1	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	0	0	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	2	2	0	0	0	0

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
6	Male	Female	MC	55	9	5	4	0	0	0
	White	Hispanic	MC	55	3	2	1	0	0	0
	White	Native American	MC	55	4	3	1	0	0	0
	No Disability	Disability	MC	55	5	4	1	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	11	6	5	11	11	0
	Male	Female	OR	5	1	0	1	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	2	2	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	1	1	0	2	1	1
7	Male	Female	MC	55	10	8	2	0	0	0
	White	Hispanic	MC	55	1	1	0	0	0	0
	White	Native American	MC	55	2	2	0	0	0	0
	No Disability	Disability	MC	55	4	2	2	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	20	12	8	6	6	0
	Male	Female	OR	5	1	0	1	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	2	2	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	0	0	0	2	2	0
8	Male	Female	MC	55	5	2	3	1	1	0
	White	Hispanic	MC	55	6	5	1	0	0	0
	White	Native American	MC	55	3	2	1	0	0	0
	No Disability	Disability	MC	55	8	7	1	1	1	0
	Not Low Income	Low Income	MC	55	1	1	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	12	9	3	11	9	2
	Male	Female	OR	5	2	0	2	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	1	1	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	2	1	1	0	0	0

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
10	Male	Female	MC	55	11	7	4	1	1	0
	White	Hispanic	MC	55	6	4	2	0	0	0
	White	Native American	MC	55	5	3	2	0	0	0
	No Disability	Disability	MC	55	17	15	2	1	1	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	55	0	0	0	0	0	0
	Male	Female	OR	5	1	0	1	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	1	1	0	0	0	0
	No Disability	Disability	OR	5	1	1	0	1	1	0
	Not Low Income	Low Income	OR	5	1	1	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	5	0	0	0	0	0	0

Table G-2. 2010–11 MontCAS: Number of Items Classified as “Low” or “High” DIF, Overall and by Group Favored – Reading

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
3	Male	Female	MC	52	0	0	0	0	0	0
	White	Hispanic	MC	52	0	0	0	1	1	0
	White	Native American	MC	52	6	6	0	0	0	0
	No Disability	Disability	MC	52	1	1	0	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	7	6	1	3	3	0
	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
4	Male	Female	MC	52	2	2	0	0	0	0
	White	Hispanic	MC	52	2	2	0	0	0	0
	White	Native American	MC	52	3	3	0	0	0	0
	No Disability	Disability	MC	52	4	4	0	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	10	8	2	8	7	1

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
4	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
5	Male	Female	MC	52	1	1	0	1	1	0
	White	Hispanic	MC	52	1	1	0	0	0	0
	White	Native American	MC	52	7	7	0	0	0	0
	No Disability	Disability	MC	52	4	4	0	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	10	8	2	6	5	1
	Male	Female	OR	2	2	0	2	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
6	Male	Female	MC	52	6	4	2	1	1	0
	White	Hispanic	MC	52	0	0	0	0	0	0
	White	Native American	MC	52	0	0	0	1	1	0
	No Disability	Disability	MC	52	5	5	0	0	0	0
	Not Low Income	Low Income	MC	52	1	1	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	12	8	4	5	4	1
	Male	Female	OR	2	2	0	2	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	1	1	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
7	Male	Female	MC	52	7	6	1	0	0	0
	White	Hispanic	MC	52	3	2	1	0	0	0
	White	Native American	MC	52	1	1	0	0	0	0
	No Disability	Disability	MC	52	1	1	0	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	15	10	5	8	7	1

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
7	Male	Female	OR	2	2	0	2	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	2	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	1	0	1	0	0	0
8	Male	Female	MC	52	8	6	2	1	1	0
	White	Hispanic	MC	52	1	1	0	0	0	0
	White	Native American	MC	52	4	3	1	0	0	0
	No Disability	Disability	MC	52	6	6	0	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	12	6	6	7	5	2
	Male	Female	OR	2	1	0	1	1	0	1
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	2	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
10	Male	Female	MC	52	4	4	0	0	0	0
	White	Hispanic	MC	52	5	5	0	0	0	0
	White	Native American	MC	52	7	7	0	0	0	0
	No Disability	Disability	MC	52	8	8	0	2	1	1
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	2	0	2	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	2	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0

**Table G-3. 2010–11 MontCAS: Number of Items Classified as “Low” or “High” DIF,
Overall and by Group Favored – Science**

Grade	Reference group	Focal group	Item type	Number of items	Number “low”			Number “high”		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
4	Male	Female	MC	53	9	5	4	1	1	0
	White	Hispanic	MC	53	1	1	0	0	0	0
	White	Native American	MC	53	6	6	0	0	0	0
	No Disability	Disability	MC	53	5	4	1	0	0	0
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	53	15	13	2	10	10	0
	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
8	Male	Female	MC	53	9	3	6	1	1	0
	White	Hispanic	MC	53	4	2	2	0	0	0
	White	Native American	MC	53	4	3	1	0	0	0
	No Disability	Disability	MC	53	13	10	3	0	0	0
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	53	15	11	4	16	12	4
	Male	Female	OR	2	1	1	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	1	1	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0
10	Male	Female	MC	53	8	5	3	1	1	0
	White	Hispanic	MC	53	8	5	3	1	1	0
	White	Native American	MC	53	4	4	0	0	0	0
	No Disability	Disability	MC	53	13	9	4	2	1	1
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	MC	53	0	0	0	0	0	0
	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	1	1	0	0	0	0
	No Disability	Disability	OR	2	2	2	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
	Not Limited English Proficient	Limited English Proficient	OR	2	0	0	0	0	0	0

Appendix H—ITEM RESPONSE THEORY CALIBRATION RESULTS

Table H-1. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 3

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
43120	1.00	-1.75842	0.00
76764	1.00	-1.09337	0.00
242748	1.00	-0.83362	0.00
138763	1.00	-0.99202	0.00
138780	1.00	-0.34528	0.00
138986	1.00	-0.0217	0.00
138787	1.00	-0.32556	0.00
138879	1.00	-0.57478	0.00
76916	1.00	0.25911	0.00
60299	1.00	-0.67421	0.00
77013	1.00	-2.073	0.00
138751	1.00	-0.82103	0.00
43112	1.00	-1.03796	0.00
138865	1.00	0.66171	0.00
76839	1.00	-0.38621	0.00
76873	1.00	-1.13416	0.00
138929	1.00	-0.92236	0.00
60375	1.00	0.07397	0.00
138756	1.00	-1.39628	0.00
138917	1.00	-0.9214	0.00
59347	1.00	-0.50965	0.00
76906	1.00	-0.44683	0.00
60271	1.00	0.92405	0.00
138982	1.00	-1.09014	0.00
138755	1.00	-0.32042	0.00
138765	1.00	-0.2651	0.00
43092	1.00	-0.24829	0.00
76968	1.00	0.10694	0.00
138832	1.00	-0.52807	0.00
42981	1.00	-1.36871	0.00
76863	1.00	0.20381	0.00
76920	1.00	0.05837	0.00
76871	1.00	-0.32497	0.00
59321	1.00	-0.49482	0.00
76780	1.00	-1.06625	0.00
138821	1.00	0.03253	0.00
34618	1.00	-0.64803	0.00
138965	1.00	-1.37099	0.00
76778	1.00	-1.04338	0.00
60276	1.00	-1.11125	0.00
76845	1.00	0.31855	0.00
60358	1.00	-0.35069	0.00
138781	1.00	-0.90157	0.00
138795	1.00	-0.45723	0.00
139031	1.00	-0.08137	0.00
138789	1.00	0.05633	0.00
206058	1.00	-1.61954	0.00
139015	1.00	-0.43813	0.00
138826	1.00	0.76182	0.00
138958	1.00	-0.03117	0.00
138995	1.00	-0.27447	0.00
76886	1.00	-0.81252	0.00
138824	1.00	0.35977	0.00
139020	1.00	-0.39745	0.00
35188	1.00	-1.6332	0.00
139045	1.00	-1.10767	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
139051	1.00	0.00529	0.00
77041	1.00	-0.11715	0.00

Table H-2. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 3

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
138799	1.00	0.19215	0.71988	-0.14358	-0.15365	-0.42265
43277	1.00	-0.01188	0.36429	0.41416	-0.05705	-0.7214

Table H-3. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 4

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
76994	1.00	-0.95182	0.00
76832	1.00	-0.55986	0.00
76789	1.00	-0.92703	0.00
77031	1.00	0.04244	0.00
77056	1.00	-0.59142	0.00
139961	1.00	-0.6836	0.00
139571	1.00	0.49582	0.00
139903	1.00	1.11441	0.00
76822	1.00	-0.42963	0.00
139538	1.00	-0.21575	0.00
139785	1.00	0.26455	0.00
77024	1.00	1.29908	0.00
62302	1.00	-0.38373	0.00
76933	1.00	0.14982	0.00
139763	1.00	0.58499	0.00
34854	1.00	-0.72639	0.00
43147	1.00	-0.41941	0.00
76804	1.00	-0.68781	0.00
76847	1.00	-0.70337	0.00
139618	1.00	-0.81686	0.00
76803	1.00	0.177	0.00
62343	1.00	-0.23361	0.00
139934	1.00	0.06595	0.00
62317	1.00	0.18406	0.00
62222	1.00	-0.03091	0.00
76891	1.00	-0.24928	0.00
77026	1.00	-0.35272	0.00
139950	1.00	0.80812	0.00
139562	1.00	-0.44442	0.00
139542	1.00	0.04929	0.00
62355	1.00	-0.80303	0.00
139592	1.00	0.00128	0.00
243107	1.00	0.24149	0.00
242919	1.00	-0.68539	0.00
248007	1.00	0.11215	0.00
77046	1.00	0.09344	0.00
43143	1.00	-0.50863	0.00
139915	1.00	-0.73775	0.00
139765	1.00	-1.07452	0.00
43266	1.00	-0.1579	0.00
62324	1.00	-0.51723	0.00
139956	1.00	0.05541	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
77022	1.00	-0.2005	0.00
248132	1.00	0.19087	0.00
77042	1.00	0.47659	0.00
76861	1.00	-0.32661	0.00
35220	1.00	-0.10912	0.00
139630	1.00	-0.41397	0.00
139545	1.00	0.13596	0.00
139911	1.00	-0.49098	0.00
34964	1.00	-0.75578	0.00
77035	1.00	0.46839	0.00
244352	1.00	-0.6317	0.00
61832	1.00	0.42944	0.00
76852	1.00	-0.86519	0.00
140171	1.00	-0.0669	0.00
76773	1.00	0.6269	0.00
76755	1.00	-0.37318	0.00

Table H-4. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 4

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
140183	1.00	0.28491	0.54666	0.32024	-0.47328	-0.39362
205515	1.00	0.29556	-0.05228	0.41674	0.76193	-1.12639

Table H-5. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 5

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
77259	1.00	-0.91828	0.00
60845	1.00	-0.58324	0.00
59861	1.00	-0.16101	0.00
43566	1.00	-0.38491	0.00
140800	1.00	0.67291	0.00
43581	1.00	-0.22527	0.00
59818	1.00	0.14981	0.00
140878	1.00	0.36392	0.00
77309	1.00	-0.33046	0.00
140823	1.00	-0.81252	0.00
77171	1.00	-0.69337	0.00
77247	1.00	-0.41913	0.00
140829	1.00	-0.52249	0.00
235941	1.00	-0.08088	0.00
77207	1.00	-0.10064	0.00
77302	1.00	-0.59257	0.00
140693	1.00	-0.07833	0.00
43437	1.00	-0.44291	0.00
43600	1.00	-0.50694	0.00
140798	1.00	-0.06743	0.00
59863	1.00	-1.09943	0.00
77254	1.00	0.61946	0.00
140907	1.00	-0.231	0.00
242883	1.00	0.45114	0.00
60383	1.00	-0.8755	0.00
43546	1.00	0.0024	0.00
77165	1.00	0.29635	0.00
43465	1.00	0.24703	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
43585	1.00	-0.31787	0.00
77205	1.00	-0.76528	0.00
77228	1.00	0.07008	0.00
43558	1.00	-0.76811	0.00
77328	1.00	-0.2042	0.00
242999	1.00	-0.68143	0.00
140786	1.00	-1.11111	0.00
43427	1.00	-0.1768	0.00
140951	1.00	-0.123	0.00
77281	1.00	-0.87545	0.00
59995	1.00	-0.59998	0.00
236242	1.00	-0.18369	0.00
59856	1.00	-0.32623	0.00
236134	1.00	0.40185	0.00
243011	1.00	-0.33792	0.00
77388	1.00	-0.66257	0.00
77169	1.00	-0.86021	0.00
77306	1.00	-0.18435	0.00
242922	1.00	0.58888	0.00
140919	1.00	0.24567	0.00
43471	1.00	-0.49379	0.00
77198	1.00	0.35794	0.00
60398	1.00	-0.33734	0.00
34525	1.00	0.27181	0.00
77183	1.00	-0.57615	0.00
77241	1.00	-0.30908	0.00
77286	1.00	-0.69558	0.00
34605	1.00	-0.4375	0.00
77299	1.00	-0.40371	0.00
62035	1.00	-0.51377	0.00

Table H-6. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 5

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
206043	1.00	-0.00621	0.8244	0.25916	-0.00418	-1.07937
206104	1.00	0.21171	0.30872	-0.04937	0.00033	-0.25967

Table H-7. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 6

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
77340	1.00	-0.12031	0.00
77377	1.00	-1.92555	0.00
140983	1.00	-0.21063	0.00
141156	1.00	0.63842	0.00
77376	1.00	-0.23274	0.00
60883	1.00	-0.10549	0.00
60890	1.00	-0.52972	0.00
43993	1.00	0.11614	0.00
77337	1.00	0.42601	0.00
141165	1.00	0.14967	0.00
43852	1.00	-1.08033	0.00
43874	1.00	0.05587	0.00
77317	1.00	0.12425	0.00
43857	1.00	0.08128	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
61130	1.00	-1.25949	0.00
43992	1.00	-0.05055	0.00
62994	1.00	0.2512	0.00
62073	1.00	-0.22893	0.00
141178	1.00	-0.08426	0.00
44040	1.00	-0.18507	0.00
61136	1.00	0.10955	0.00
77335	1.00	0.13288	0.00
141406	1.00	-0.62593	0.00
62060	1.00	-0.94161	0.00
61162	1.00	-0.21394	0.00
141175	1.00	-0.0428	0.00
44027	1.00	0.37807	0.00
62046	1.00	-0.56801	0.00
77577	1.00	-0.22633	0.00
61151	1.00	-0.1338	0.00
77540	1.00	-0.11909	0.00
141476	1.00	-0.50223	0.00
77414	1.00	-0.6326	0.00
141292	1.00	0.09768	0.00
141272	1.00	1.53414	0.00
77621	1.00	-0.56638	0.00
77531	1.00	-0.47197	0.00
141405	1.00	0.17354	0.00
77461	1.00	0.03224	0.00
141452	1.00	-0.29636	0.00
77339	1.00	0.078	0.00
141283	1.00	-0.79446	0.00
243207	1.00	0.01726	0.00
43985	1.00	-0.48645	0.00
62047	1.00	0.11308	0.00
43944	1.00	-0.20245	0.00
77555	1.00	-0.77871	0.00
62044	1.00	-0.10672	0.00
141434	1.00	-0.16222	0.00
141470	1.00	-0.27242	0.00
141407	1.00	-0.59034	0.00
140815	1.00	0.52655	0.00
62062	1.00	-0.6103	0.00
59842	1.00	0.04883	0.00
62054	1.00	-0.57659	0.00
141512	1.00	-0.4327	0.00
77646	1.00	-0.56341	0.00
43913	1.00	-0.07634	0.00

Table H-8. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 6

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
63054	1.00	0.31923	0.91617	-0.06101	-0.92973	0.07456
206762	1.00	0.75164	0.49972	0.31367	-0.25397	-0.55942

Table H-9. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 7

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
86269	1.00	-0.36486	0.00
142365	1.00	-0.40164	0.00
86290	1.00	-0.81099	0.00
43791	1.00	0.21211	0.00
86288	1.00	-0.16795	0.00
77389	1.00	0.40746	0.00
142368	1.00	-0.30405	0.00
86264	1.00	0.22597	0.00
86330	1.00	-0.17773	0.00
86273	1.00	-0.09396	0.00
86302	1.00	0.01231	0.00
43836	1.00	-0.26535	0.00
43862	1.00	0.50268	0.00
43842	1.00	-1.22564	0.00
61766	1.00	-1.11401	0.00
86624	1.00	-0.49609	0.00
142680	1.00	0.61301	0.00
61244	1.00	-0.8882	0.00
86382	1.00	-0.71731	0.00
142645	1.00	-0.57446	0.00
142735	1.00	0.73333	0.00
142377	1.00	0.49763	0.00
86549	1.00	0.03723	0.00
86606	1.00	0.02725	0.00
86687	1.00	0.22366	0.00
86672	1.00	-0.44391	0.00
142376	1.00	-0.06978	0.00
61781	1.00	-0.62468	0.00
61363	1.00	-0.03015	0.00
86597	1.00	0.20893	0.00
43654	1.00	-0.45415	0.00
86568	1.00	0.09544	0.00
61279	1.00	-1.20479	0.00
43753	1.00	-0.17651	0.00
86431	1.00	0.33406	0.00
86535	1.00	-0.84153	0.00
86700	1.00	0.85358	0.00
61799	1.00	-0.25243	0.00
61874	1.00	-0.22432	0.00
86681	1.00	-0.82845	0.00
86445	1.00	0.42888	0.00
43680	1.00	0.61054	0.00
86333	1.00	-0.19508	0.00
43767	1.00	-1.26714	0.00
142818	1.00	0.32061	0.00
61745	1.00	-0.29149	0.00
142649	1.00	-0.29441	0.00
44173	1.00	0.78686	0.00
61348	1.00	0.02722	0.00
86644	1.00	-0.15675	0.00
61772	1.00	-1.30768	0.00
86574	1.00	-0.1432	0.00
86641	1.00	-0.18838	0.00
43780	1.00	-0.41104	0.00
43719	1.00	-0.08218	0.00
43911	1.00	0.04303	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
61374	1.00	-0.42905	0.00
86339	1.00	0.11454	0.00

Table H-10. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 7

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
142673	1.00	0.63551	0.91044	0.40162	-0.47991	-0.83215
43828	1.00	0.10673	0.29804	0.89119	0.21389	-1.40312

Table H-11. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 8

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
144251	1.00	-0.35428	0.00
144191	1.00	-0.31625	0.00
87653	1.00	-0.01686	0.00
87576	1.00	0.18424	0.00
87566	1.00	0.55731	0.00
88026	1.00	-0.65459	0.00
87571	1.00	-0.07234	0.00
43840	1.00	0.53935	0.00
35025	1.00	-0.44952	0.00
44181	1.00	0.06732	0.00
63242	1.00	-0.24131	0.00
61198	1.00	0.37852	0.00
244462	1.00	0.14559	0.00
44209	1.00	-0.38185	0.00
88177	1.00	-0.57249	0.00
44127	1.00	-0.74114	0.00
244461	1.00	-0.27717	0.00
244684	1.00	0.20067	0.00
35019	1.00	-1.17433	0.00
44213	1.00	0.70974	0.00
43888	1.00	0.0351	0.00
144456	1.00	-0.48112	0.00
144499	1.00	-0.34305	0.00
63048	1.00	-0.60857	0.00
63239	1.00	-0.02959	0.00
88892	1.00	0.02723	0.00
88277	1.00	0.33189	0.00
63109	1.00	0.5956	0.00
88838	1.00	-0.33409	0.00
166323	1.00	-1.00E-05	0.00
144905	1.00	-0.68256	0.00
34969	1.00	0.98273	0.00
34932	1.00	0.60652	0.00
88352	1.00	-0.36843	0.00
88216	1.00	-0.42302	0.00
34995	1.00	-1.28055	0.00
63038	1.00	-0.55482	0.00
144955	1.00	-0.3931	0.00
144942	1.00	-0.1033	0.00
144436	1.00	0.21714	0.00
34982	1.00	-0.24118	0.00
88361	1.00	0.00685	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
63115	1.00	0.2962	0.00
44210	1.00	-0.25901	0.00
88313	1.00	0.0045	0.00
88086	1.00	0.06331	0.00
63066	1.00	0.43647	0.00
86542	1.00	-0.54125	0.00
44175	1.00	0.74157	0.00
44154	1.00	0.04347	0.00
88187	1.00	0.51647	0.00
88331	1.00	0.18751	0.00
34933	1.00	0.03346	0.00
63256	1.00	-0.43442	0.00
63106	1.00	-0.86658	0.00
243715	1.00	0.04168	0.00
243699	1.00	0.03673	0.00
243710	1.00	0.35706	0.00

Table H-12. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 8

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
87874	1.00	0.08265	-0.00596	1.02254	-0.56458	-0.452
205490	1.00	0.15854	0.80473	-0.0317	-0.81978	0.04674

Table H-13. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 10

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
146070	1.00	-0.08126	0.00
77592	1.00	0.27254	0.00
77369	1.00	-0.22723	0.00
144377	1.00	-0.23806	0.00
43725	1.00	0.74114	0.00
77512	1.00	0.77145	0.00
77364	1.00	-1.03123	0.00
59403	1.00	-0.32718	0.00
144570	1.00	0.19984	0.00
77501	1.00	0.39321	0.00
145022	1.00	-0.01269	0.00
59371	1.00	0.79473	0.00
34804	1.00	0.04766	0.00
43687	1.00	0.43715	0.00
62361	1.00	-0.14499	0.00
62279	1.00	-0.4342	0.00
144996	1.00	0.07827	0.00
145032	1.00	0.13411	0.00
146556	1.00	0.24026	0.00
146573	1.00	0.48533	0.00
43951	1.00	0.586	0.00
77546	1.00	0.22241	0.00
145719	1.00	0.10364	0.00
144825	1.00	-0.09444	0.00
145637	1.00	-0.18966	0.00
77428	1.00	-0.27176	0.00
77374	1.00	0.6513	0.00
77602	1.00	0.93806	0.00
146535	1.00	-0.00373	0.00

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
77358	1.00	-0.21413	0.00
146539	1.00	-0.04584	0.00
77420	1.00	0.54531	0.00
77424	1.00	0.86044	0.00
144896	1.00	-0.00971	0.00
62380	1.00	-0.00834	0.00
145292	1.00	-0.10928	0.00
43969	1.00	-0.98259	0.00
61301	1.00	0.03673	0.00
77430	1.00	0.05609	0.00
144841	1.00	-0.02886	0.00
43717	1.00	-0.60682	0.00
77544	1.00	-0.07344	0.00
243050	1.00	-0.30551	0.00
146572	1.00	0.00147	0.00
43631	1.00	0.75885	0.00
146580	1.00	0.04044	0.00
248806	1.00	0.49642	0.00
61324	1.00	-1.31138	0.00
62319	1.00	0.15909	0.00
61305	1.00	0.72894	0.00
62365	1.00	0.21208	0.00
144904	1.00	-0.54862	0.00
77435	1.00	1.70657	0.00
77503	1.00	0.59545	0.00
77432	1.00	-0.29149	0.00
77644	1.00	0.89482	0.00
145015	1.00	-0.41016	0.00
77629	1.00	0.25914	0.00

Table H-14. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 10

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
62391	1.00	0.76356	0.46836	0.22485	-0.27884	-0.41438
59407	1.00	0.40229	0.35406	0.07648	-0.72152	0.29098

Table H-15. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 3

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
150711	1.00	-0.1012	0.00
150713	1.00	-0.04271	0.00
150718	1.00	-0.99896	0.00
150721	1.00	-1.18962	0.00
150725	1.00	-1.50459	0.00
150869	1.00	-1.42395	0.00
150951	1.00	-1.07128	0.00
151173	1.00	-0.45488	0.00
151176	1.00	-1.04983	0.00
151186	1.00	-0.54915	0.00
151193	1.00	0.00379	0.00
151194	1.00	0.21961	0.00
151200	1.00	-0.71684	0.00
151203	1.00	-0.35603	0.00
151207	1.00	0.38192	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
151212	1.00	-0.1503	0.00
151215	1.00	-0.75997	0.00
151227	1.00	-0.05165	0.00
153156	1.00	-0.12623	0.00
67148	1.00	-0.46242	0.00
67151	1.00	-0.51294	0.00
67155	1.00	-1.26487	0.00
67167	1.00	-0.86298	0.00
67184	1.00	-0.44205	0.00
67193	1.00	-0.59805	0.00
67198	1.00	0.19875	0.00
151127	1.00	-1.35325	0.00
151131	1.00	-0.17644	0.00
151133	1.00	-0.57426	0.00
151141	1.00	-0.02362	0.00
151135	1.00	-0.51942	0.00
151156	1.00	-0.83513	0.00
151166	1.00	0.09425	0.00
150634	1.00	0.37209	0.00
150644	1.00	-0.02464	0.00
150648	1.00	-0.92585	0.00
150653	1.00	-1.29492	0.00
150656	1.00	-0.06756	0.00
150664	1.00	-0.98738	0.00
150670	1.00	-0.69148	0.00
150953	1.00	-0.51644	0.00
150962	1.00	-0.91022	0.00
150969	1.00	-0.721	0.00
150971	1.00	-1.46356	0.00
150974	1.00	-0.54248	0.00
150975	1.00	-1.50711	0.00
150977	1.00	-0.50999	0.00
150978	1.00	-1.44418	0.00
150982	1.00	0.30049	0.00
150983	1.00	-0.84772	0.00
150986	1.00	-0.35316	0.00
150988	1.00	0.06329	0.00

Table H-16. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 3

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
151240	1.00	0.58162	1.13818	0.04655	-0.51226	-0.67247
150991	1.00	0.28425	1.04315	0.38337	-0.47092	-0.9556

Table H-17. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 4

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
67140	1.00	-0.19441	0.00
67149	1.00	-0.74181	0.00
67154	1.00	-0.78611	0.00
67158	1.00	0.11123	0.00
67164	1.00	-0.43327	0.00
67160	1.00	-0.57309	0.00
67166	1.00	-0.61781	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
151672	1.00	-0.7856	0.00
151679	1.00	0.22177	0.00
151683	1.00	-0.56501	0.00
151684	1.00	0.15188	0.00
151686	1.00	-0.0693	0.00
151695	1.00	-0.34982	0.00
151697	1.00	-0.55278	0.00
151740	1.00	-0.22843	0.00
151743	1.00	-0.86147	0.00
151745	1.00	-0.10871	0.00
151751	1.00	-0.27758	0.00
151762	1.00	0.07422	0.00
151780	1.00	0.72873	0.00
151782	1.00	-0.90798	0.00
151783	1.00	-0.86182	0.00
151788	1.00	-0.07903	0.00
151790	1.00	-0.23214	0.00
151794	1.00	0.03287	0.00
151800	1.00	-0.48099	0.00
66546	1.00	0.25879	0.00
66550	1.00	-0.67793	0.00
66555	1.00	0.18314	0.00
66568	1.00	-0.40538	0.00
66571	1.00	-0.26656	0.00
66578	1.00	0.39532	0.00
66586	1.00	-0.22577	0.00
66597	1.00	-0.70036	0.00
66612	1.00	-0.20321	0.00
66650	1.00	-0.73576	0.00
66615	1.00	-0.91476	0.00
66629	1.00	-0.33449	0.00
66627	1.00	-0.66408	0.00
66634	1.00	-0.06523	0.00
151597	1.00	-0.12243	0.00
151612	1.00	-0.60555	0.00
151615	1.00	-0.70024	0.00
151621	1.00	0.07058	0.00
151626	1.00	-0.70785	0.00
151632	1.00	0.42225	0.00
151635	1.00	-0.13115	0.00
151637	1.00	-0.75604	0.00
151655	1.00	0.1941	0.00
151638	1.00	-0.09943	0.00
151639	1.00	0.0125	0.00
151644	1.00	-0.20892	0.00

Table H-18. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 4

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
151772	1.00	0.6713	1.67954	0.27673	-0.78618	-1.17009
151659	1.00	0.86022	1.25682	0.20629	-0.60095	-0.86216

Table H-19. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 5

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>C</i>
150580	1.00	-0.97903	0.00
150578	1.00	-0.15621	0.00
150586	1.00	-0.426	0.00
150591	1.00	-0.78255	0.00
150593	1.00	-0.09947	0.00
150596	1.00	-0.05856	0.00
150594	1.00	-0.48815	0.00
150415	1.00	-0.76413	0.00
150413	1.00	-0.02298	0.00
150417	1.00	-0.74223	0.00
150419	1.00	-0.45492	0.00
200624	1.00	-0.58895	0.00
150425	1.00	-0.47642	0.00
150427	1.00	-1.0139	0.00
150423	1.00	-0.38503	0.00
150432	1.00	0.28726	0.00
150440	1.00	-1.1782	0.00
150442	1.00	-0.7183	0.00
150450	1.00	0.6174	0.00
65580	1.00	0.10519	0.00
65593	1.00	-0.76817	0.00
65616	1.00	-0.85416	0.00
65599	1.00	-0.18149	0.00
65607	1.00	-0.3002	0.00
65611	1.00	0.12799	0.00
65613	1.00	0.05247	0.00
155431	1.00	-0.51061	0.00
150527	1.00	-0.024	0.00
150536	1.00	-0.36111	0.00
150530	1.00	-0.94657	0.00
150547	1.00	-0.78942	0.00
150548	1.00	0.15085	0.00
150551	1.00	-0.49183	0.00
41321	1.00	-0.54104	0.00
41322	1.00	0.32399	0.00
41323	1.00	-0.5404	0.00
41326	1.00	-0.80477	0.00
41328	1.00	-0.3303	0.00
41329	1.00	-0.16283	0.00
41331	1.00	-0.88234	0.00
150470	1.00	-1.19564	0.00
150480	1.00	-1.0046	0.00
150474	1.00	-0.52933	0.00
150483	1.00	-0.65197	0.00
150485	1.00	-0.5307	0.00
150488	1.00	-1.10399	0.00
150489	1.00	-0.24771	0.00
150491	1.00	-0.73095	0.00
150492	1.00	-0.76447	0.00
150493	1.00	-0.69456	0.00
150494	1.00	-0.60482	0.00
150505	1.00	-0.1321	0.00

Table H-20. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 5

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
150456	1.00	0.39017	1.60803	0.10694	-0.60937	-1.1056
150518	1.00	0.18598	1.65961	0.12246	-0.56198	-1.2201

Table H-21. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 6

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
42118	1.00	-1.06601	0.00
42131	1.00	-0.89924	0.00
42130	1.00	-1.08238	0.00
42282	1.00	-0.37038	0.00
42132	1.00	-0.82294	0.00
42287	1.00	-1.02684	0.00
42284	1.00	-0.35653	0.00
95414	1.00	-0.78695	0.00
95417	1.00	-0.17541	0.00
95422	1.00	-0.71083	0.00
95426	1.00	0.42935	0.00
95447	1.00	-0.98632	0.00
95429	1.00	-1.02101	0.00
95446	1.00	-0.23985	0.00
95456	1.00	0.25171	0.00
95448	1.00	-0.56174	0.00
95451	1.00	0.01399	0.00
95459	1.00	-0.29052	0.00
95455	1.00	-0.39742	0.00
151296	1.00	-0.82149	0.00
151295	1.00	-0.35946	0.00
151278	1.00	-0.51434	0.00
151274	1.00	-0.21205	0.00
151273	1.00	-0.98499	0.00
151282	1.00	-0.7631	0.00
151302	1.00	0.29967	0.00
151312	1.00	-0.05833	0.00
151315	1.00	-0.37777	0.00
151321	1.00	-0.11923	0.00
151325	1.00	-0.10176	0.00
151337	1.00	-0.7534	0.00
151338	1.00	0.28485	0.00
151341	1.00	-0.17306	0.00
42100	1.00	-1.11051	0.00
42099	1.00	-0.49909	0.00
42102	1.00	-0.59548	0.00
201077	1.00	-0.77563	0.00
42104	1.00	-1.0843	0.00
201078	1.00	-0.30882	0.00
42106	1.00	0.16073	0.00
68240	1.00	0.17432	0.00
68242	1.00	-0.16393	0.00
68247	1.00	-0.33248	0.00
68250	1.00	0.25394	0.00
68260	1.00	-0.37059	0.00
68263	1.00	-0.15889	0.00
68291	1.00	-0.03904	0.00
68283	1.00	-0.54226	0.00
201084	1.00	0.15409	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
68265	1.00	-0.4231	0.00
68285	1.00	-0.26609	0.00
68269	1.00	0.25911	0.00

Table H-22. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 6

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
95469	1.00	0.45974	1.34933	0.35101	-0.57886	-1.12148
68294	1.00	0.621	1.3941	0.22471	-0.46379	-1.15501

Table H-23. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 7

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
149061	1.00	-0.39141	0.00
149064	1.00	-0.7605	0.00
149066	1.00	-0.5658	0.00
154487	1.00	-0.85249	0.00
200972	1.00	-0.8517	0.00
200973	1.00	-0.19485	0.00
149080	1.00	-0.50696	0.00
148792	1.00	-0.50056	0.00
148807	1.00	-0.16955	0.00
148811	1.00	-0.0538	0.00
148827	1.00	-0.16132	0.00
148834	1.00	-0.76986	0.00
148839	1.00	-0.96219	0.00
148843	1.00	-0.30331	0.00
148849	1.00	-0.4668	0.00
148854	1.00	-0.57327	0.00
200974	1.00	-0.73091	0.00
148875	1.00	-0.55995	0.00
148883	1.00	-0.44765	0.00
92453	1.00	0.14146	0.00
92455	1.00	-1.03076	0.00
92462	1.00	-0.29222	0.00
92464	1.00	-0.57644	0.00
92458	1.00	-0.39421	0.00
92471	1.00	-0.78696	0.00
92472	1.00	-0.40686	0.00
92359	1.00	-0.64772	0.00
92384	1.00	-1.44606	0.00
92399	1.00	-0.43919	0.00
92391	1.00	-1.07368	0.00
92395	1.00	-0.38642	0.00
92397	1.00	-0.96949	0.00
92404	1.00	-0.71934	0.00
149083	1.00	-0.47476	0.00
149088	1.00	-0.49815	0.00
149089	1.00	-0.72216	0.00
149090	1.00	-0.4924	0.00
149091	1.00	-0.71616	0.00
149093	1.00	0.12549	0.00
149098	1.00	-0.01351	0.00
148965	1.00	-0.58088	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
200975	1.00	-0.30225	0.00
148970	1.00	-0.61471	0.00
148972	1.00	-0.8807	0.00
148973	1.00	-0.96704	0.00
148977	1.00	-0.20067	0.00
148982	1.00	0.04505	0.00
148988	1.00	-0.74326	0.00
148993	1.00	0.27246	0.00
149003	1.00	-0.53158	0.00
149007	1.00	-0.77447	0.00
149008	1.00	-0.1471	0.00

Table H-24. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 7

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
148887	1.00	0.24916	1.03033	0.29262	-0.45518	-0.86776
149016	1.00	0.32246	1.01897	0.28419	-0.37651	-0.92664

Table H-25. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 8

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
149129	1.00	0.0061	0.00
149130	1.00	0.37185	0.00
149134	1.00	-0.14913	0.00
149136	1.00	0.15948	0.00
149139	1.00	-0.28901	0.00
152410	1.00	-0.31134	0.00
149144	1.00	-0.29175	0.00
149342	1.00	0.27617	0.00
149344	1.00	-0.22773	0.00
149347	1.00	0.08236	0.00
152841	1.00	-0.2445	0.00
149349	1.00	0.19012	0.00
149352	1.00	-0.35263	0.00
149353	1.00	-0.25062	0.00
149354	1.00	0.02985	0.00
149355	1.00	0.10773	0.00
149360	1.00	-0.2791	0.00
149359	1.00	-0.32972	0.00
149362	1.00	-0.84701	0.00
149326	1.00	-0.05424	0.00
149328	1.00	-0.72005	0.00
149331	1.00	-0.21131	0.00
149339	1.00	-0.71227	0.00
149340	1.00	-0.16877	0.00
149341	1.00	0.00619	0.00
149335	1.00	-1.13776	0.00
153228	1.00	-0.57308	0.00
149261	1.00	-0.18653	0.00
149262	1.00	-0.00357	0.00
149269	1.00	-0.66804	0.00
149276	1.00	-0.13068	0.00
149281	1.00	0.12556	0.00
149287	1.00	0.02077	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
149372	1.00	0.45421	0.00
149388	1.00	-0.02209	0.00
153158	1.00	0.01517	0.00
149377	1.00	-0.02152	0.00
149379	1.00	-0.4205	0.00
149380	1.00	-0.53857	0.00
149383	1.00	-0.17367	0.00
149174	1.00	0.22596	0.00
149188	1.00	-0.32557	0.00
149165	1.00	0.06298	0.00
149170	1.00	-0.67401	0.00
149171	1.00	-0.73805	0.00
156089	1.00	0.16265	0.00
149176	1.00	-0.19897	0.00
149179	1.00	-0.75199	0.00
149181	1.00	-0.31959	0.00
149187	1.00	-0.38923	0.00
149190	1.00	-0.17456	0.00
149191	1.00	-0.42021	0.00

Table H-26. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 8

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
149368	1.00	0.36798	0.97942	0.35305	-0.28944	-1.04303
149193	1.00	0.31154	1.02786	0.63317	-0.35059	-1.31045

Table H-27. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 10

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
149467	1.00	-0.57102	0.00
149468	1.00	0.09244	0.00
149471	1.00	0.01215	0.00
149472	1.00	-0.11485	0.00
149482	1.00	-0.6413	0.00
149483	1.00	-0.17378	0.00
149486	1.00	-0.62829	0.00
149545	1.00	-0.1753	0.00
149548	1.00	-0.57784	0.00
149558	1.00	-0.05013	0.00
149544	1.00	-0.66591	0.00
149551	1.00	0.24022	0.00
149550	1.00	-0.15268	0.00
149549	1.00	-0.16083	0.00
149560	1.00	-0.00582	0.00
149555	1.00	-0.5683	0.00
149554	1.00	-0.29364	0.00
149556	1.00	-0.58234	0.00
149564	1.00	-0.18718	0.00
42224	1.00	-0.24508	0.00
42225	1.00	-0.27199	0.00
42226	1.00	-0.38321	0.00
42228	1.00	-0.32053	0.00
42230	1.00	-0.02617	0.00
42233	1.00	-0.32053	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
42232	1.00	0.00058	0.00
149894	1.00	-0.19715	0.00
149896	1.00	-0.25617	0.00
149869	1.00	0.13035	0.00
149883	1.00	-0.67761	0.00
149887	1.00	-0.11535	0.00
149905	1.00	-0.38526	0.00
149950	1.00	-1.0768	0.00
149515	1.00	-0.09898	0.00
149521	1.00	-0.26902	0.00
149525	1.00	-0.34203	0.00
149528	1.00	-0.37318	0.00
149538	1.00	-0.37618	0.00
149537	1.00	-0.75122	0.00
149541	1.00	-0.78704	0.00
149610	1.00	0.18182	0.00
149611	1.00	-0.70308	0.00
149613	1.00	0.42392	0.00
149615	1.00	-0.3459	0.00
149616	1.00	0.29104	0.00
149627	1.00	-0.66966	0.00
149633	1.00	-0.18153	0.00
149623	1.00	0.42738	0.00
149629	1.00	-0.31766	0.00
149638	1.00	0.04269	0.00
149630	1.00	0.3285	0.00
149631	1.00	-0.26813	0.00

Table H-28. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 10

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
149566	1.00	0.58296	0.93261	0.41333	-0.47282	-0.87312
149648	1.00	0.57591	0.63281	0.39838	-0.3226	-0.70859

Table H-29. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 4

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
39203	1.00	-0.17019	0.00
120003	1.00	-0.36305	0.00
75718	1.00	-0.55467	0.00
75719	1.00	-0.10849	0.00
194487	1.00	-1.58446	0.00
120024	1.00	0.1006	0.00
38537	1.00	-0.73275	0.00
119986	1.00	-0.55613	0.00
75818	1.00	-1.50675	0.00
120106	1.00	0.3085	0.00
76397	1.00	0.32507	0.00
142358	1.00	0.04957	0.00
134838	1.00	-0.4847	0.00
39164	1.00	-0.72509	0.00
75708	1.00	-0.60778	0.00
76401	1.00	-0.23429	0.00
75831	1.00	-0.9835	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
134736	1.00	-0.54719	0.00
134742	1.00	-1.61346	0.00
120564	1.00	-1.54291	0.00
120617	1.00	-0.95381	0.00
119971	1.00	-1.08833	0.00
134858	1.00	-0.81907	0.00
52579	1.00	-1.29519	0.00
119973	1.00	-0.22089	0.00
57902	1.00	-0.52209	0.00
75760	1.00	-0.80911	0.00
75691	1.00	-1.27725	0.00
76289	1.00	-0.80529	0.00
75692	1.00	-0.15205	0.00
56263	1.00	0.07271	0.00
144837	1.00	0.00191	0.00
53280	1.00	-1.01383	0.00
75693	1.00	-0.73542	0.00
57920	1.00	0.02907	0.00
75774	1.00	-0.04726	0.00
75695	1.00	-1.33998	0.00
120543	1.00	-0.20647	0.00
120162	1.00	-0.35717	0.00
134745	1.00	0.16335	0.00
75424	1.00	-0.40228	0.00
120014	1.00	-0.53667	0.00
75406	1.00	-0.26344	0.00
75513	1.00	-0.58213	0.00
134675	1.00	-0.37925	0.00
55442	1.00	-0.06762	0.00
56367	1.00	-0.11747	0.00
75405	1.00	-0.29095	0.00
56225	1.00	0.02096	0.00
56327	1.00	-0.1487	0.00
75412	1.00	-0.53918	0.00
75404	1.00	-0.56658	0.00
75399	1.00	0.36753	0.00

Table H-30. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Science Grade 4

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
75757	1.00	-0.31094	0.6748	-0.01752	-0.48603	-0.17125
120312	1.00	0.04833	0.7182	0.89294	-0.75841	-0.85273

Table H-31. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 8

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
89277	1.00	-1.01669	0.00
89752	1.00	-0.41017	0.00
89591	1.00	-0.09389	0.00
54228	1.00	-1.2942	0.00
89859	1.00	0.2284	0.00
122023	1.00	-0.24867	0.00
121232	1.00	-0.16977	0.00
89392	1.00	0.32839	0.00

continued

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
89361	1.00	-0.38364	0.00
39587	1.00	-1.58575	0.00
89900	1.00	-0.2353	0.00
56879	1.00	0.19626	0.00
144065	1.00	-0.64154	0.00
56773	1.00	-0.74439	0.00
89504	1.00	-0.50224	0.00
89892	1.00	-0.32073	0.00
56876	1.00	-0.04656	0.00
89433	1.00	0.10313	0.00
56989	1.00	0.41284	0.00
89644	1.00	0.13703	0.00
39701	1.00	-0.56951	0.00
122738	1.00	0.04677	0.00
89860	1.00	2.7669	0.00
56846	1.00	-0.4015	0.00
89637	1.00	0.01907	0.00
89917	1.00	-1.00789	0.00
89934	1.00	-0.05612	0.00
39729	1.00	0.20897	0.00
39773	1.00	0.02752	0.00
89895	1.00	-0.96498	0.00
122736	1.00	-0.46944	0.00
89702	1.00	-0.22607	0.00
39648	1.00	-0.77746	0.00
89426	1.00	-0.32411	0.00
89901	1.00	0.25328	0.00
89857	1.00	-0.39221	0.00
121209	1.00	0.26071	0.00
89621	1.00	0.07524	0.00
89406	1.00	-0.53157	0.00
56985	1.00	0.13445	0.00
54237	1.00	-0.87022	0.00
89463	1.00	0.61301	0.00
39782	1.00	0.00803	0.00
56894	1.00	-0.82032	0.00
89710	1.00	0.3698	0.00
89881	1.00	-0.1831	0.00
121222	1.00	-0.4373	0.00
122036	1.00	-1.01506	0.00
39780	1.00	-0.75313	0.00
89914	1.00	0.26338	0.00
89513	1.00	-1.05606	0.00
39546	1.00	0.99728	0.00
89766	1.00	-0.33793	0.00

Table H-32. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Science Grade 8

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
89534	1.00	0.56178	0.73836	0.09856	-0.16965	-0.66728
89958	1.00	0.32235	0.40732	0.27961	-0.19758	-0.48935

Table H-33. 2010–11 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 10

<i>IREF</i>	<i>Parameters</i>		
	<i>a</i>	<i>b</i>	<i>c</i>
134543	1.00	-0.2538	0.00
134489	1.00	-0.622	0.00
197896	1.00	-0.30089	0.00
119676	1.00	0.21336	0.00
75644	1.00	0.34927	0.00
75433	1.00	-0.05999	0.00
75441	1.00	-0.09862	0.00
134482	1.00	-0.52609	0.00
75865	1.00	-0.00674	0.00
120311	1.00	-0.37894	0.00
119661	1.00	-0.21327	0.00
130550	1.00	-1.22923	0.00
134545	1.00	-0.28319	0.00
75786	1.00	-0.53125	0.00
40188	1.00	0.51161	0.00
134487	1.00	-0.92642	0.00
75976	1.00	-0.80404	0.00
75963	1.00	-0.24741	0.00
134480	1.00	-0.07654	0.00
75875	1.00	-0.96624	0.00
119955	1.00	0.04115	0.00
120064	1.00	-0.83395	0.00
75859	1.00	-0.20733	0.00
134799	1.00	-0.80975	0.00
75635	1.00	-0.01941	0.00
75941	1.00	-0.46638	0.00
75725	1.00	0.43249	0.00
134778	1.00	0.25969	0.00
119674	1.00	-0.14317	0.00
75803	1.00	-0.16251	0.00
134770	1.00	-0.26199	0.00
119664	1.00	0.35289	0.00
75452	1.00	0.04446	0.00
75632	1.00	0.07224	0.00
52960	1.00	0.21041	0.00
119943	1.00	-0.51802	0.00
75787	1.00	-0.966	0.00
134792	1.00	-0.49713	0.00
75620	1.00	0.45165	0.00
75804	1.00	0.09878	0.00
75726	1.00	0.65287	0.00
134479	1.00	-0.1337	0.00
119675	1.00	0.25969	0.00
130594	1.00	-0.15583	0.00
75863	1.00	-0.4104	0.00
130592	1.00	-0.2523	0.00
75634	1.00	0.1591	0.00
75854	1.00	-0.21561	0.00
119656	1.00	-0.22944	0.00
75630	1.00	0.1998	0.00
75970	1.00	-0.13565	0.00
75777	1.00	-0.76284	0.00
75846	1.00	-0.59992	0.00

Table H-34. 2010–11 MontCAS: IRT Parameters for Polytomous Items – Science Grade 10

<i>IREF</i>	<i>Parameters</i>					
	<i>a</i>	<i>b</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>
119960	1.00	0.45817	0.5493	0.51508	-0.47716	-0.58722
52953	1.00	0.65676	1.32929	0.13605	-0.19433	-1.27101

Appendix I—TCCs AND TIFs

Figure I-1. 2010–11 MontCAS: TCC – Mathematics Grade 3

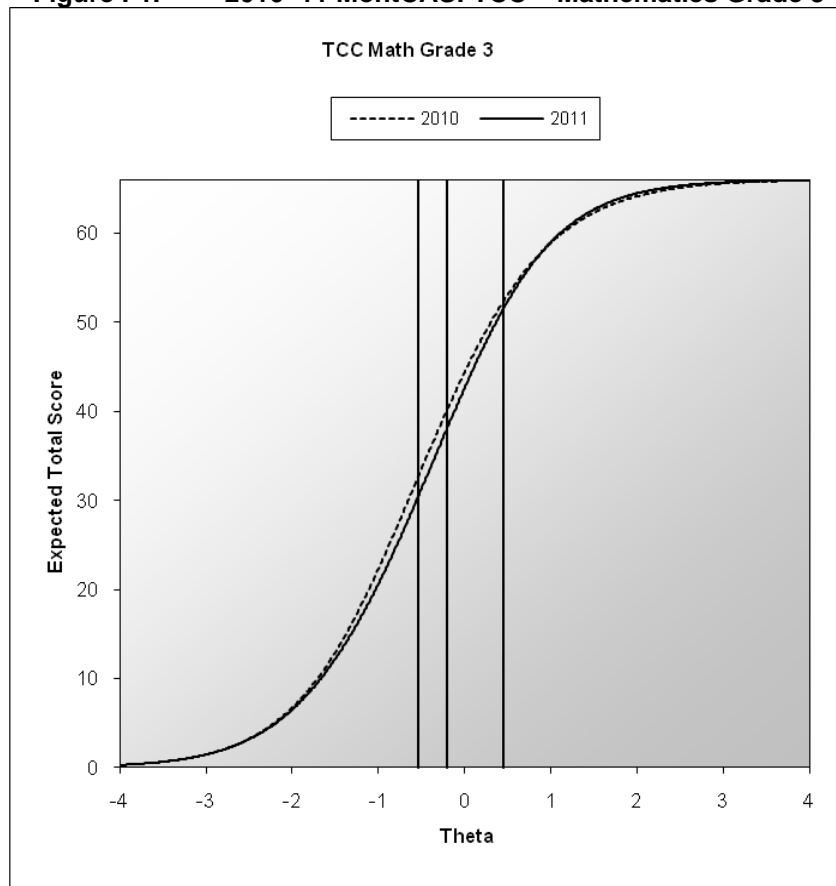


Figure I-2. 2010–11 MontCAS: TIF – Mathematics Grade 3

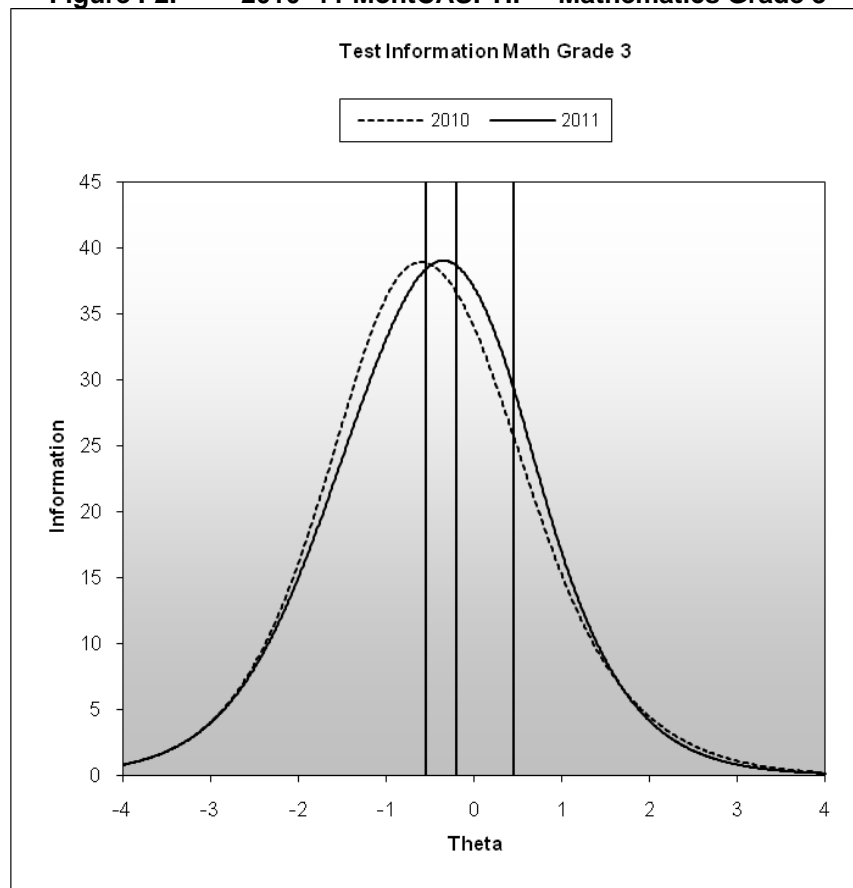


Figure I-3. 2010–11 MontCAS: TCC – Mathematics Grade 4

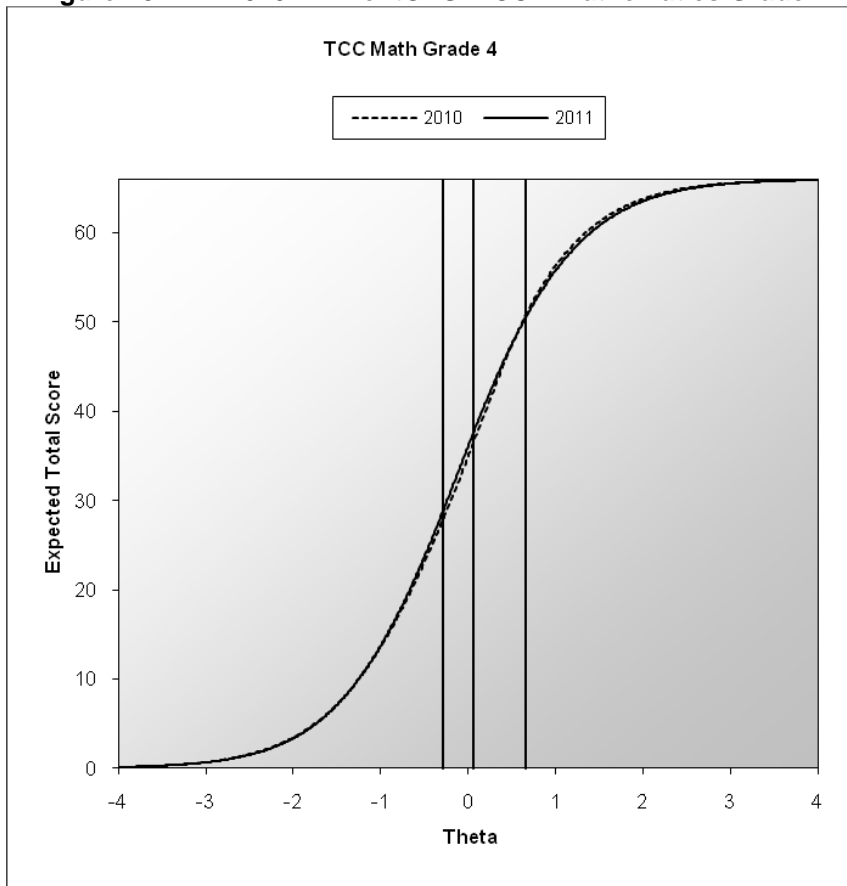


Figure I-4. 2010–11 MontCAS: TIF – Mathematics Grade 4

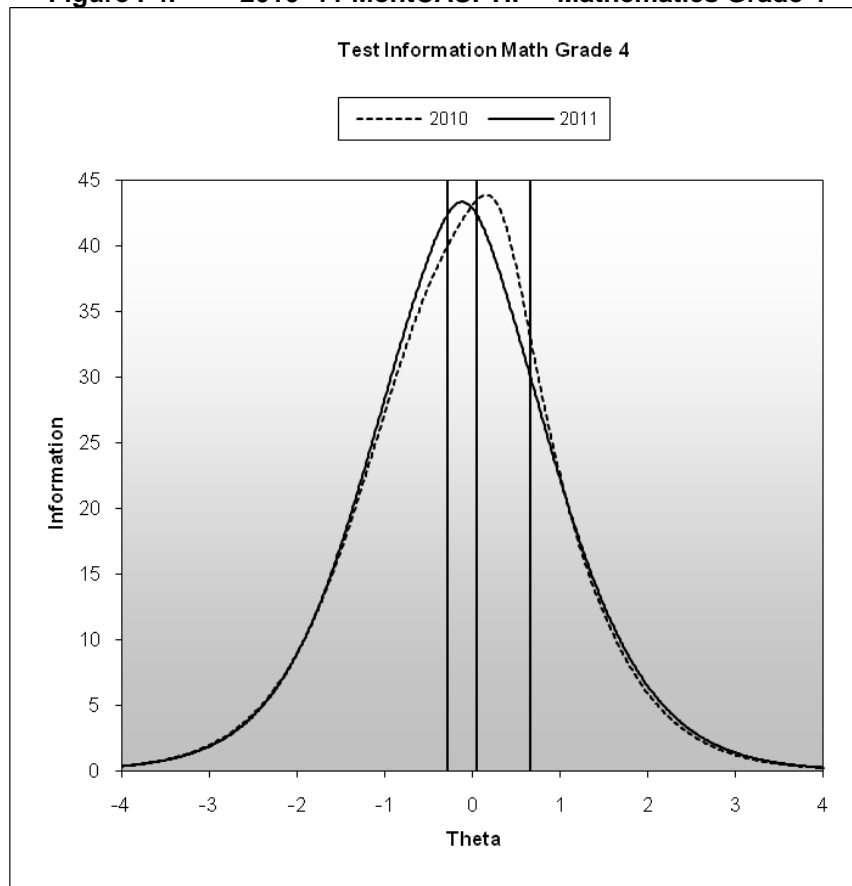


Figure I-5. 2010–11 MontCAS: TCC – Mathematics Grade 5

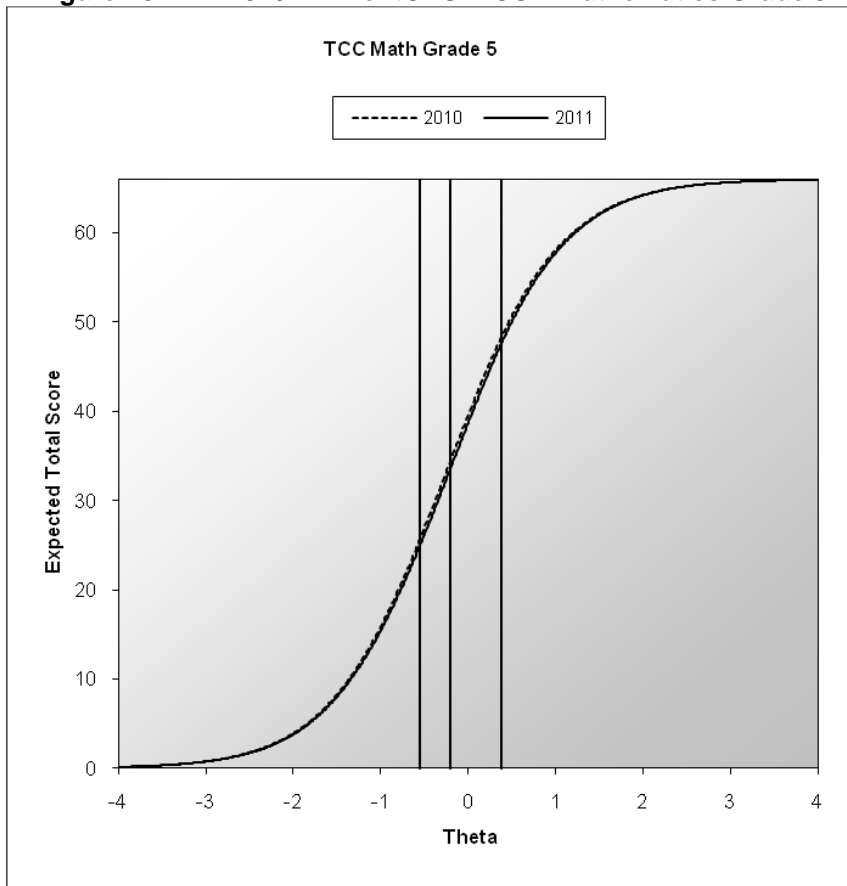


Figure I-6. 2010–11 MontCAS: TIF – Mathematics Grade 5

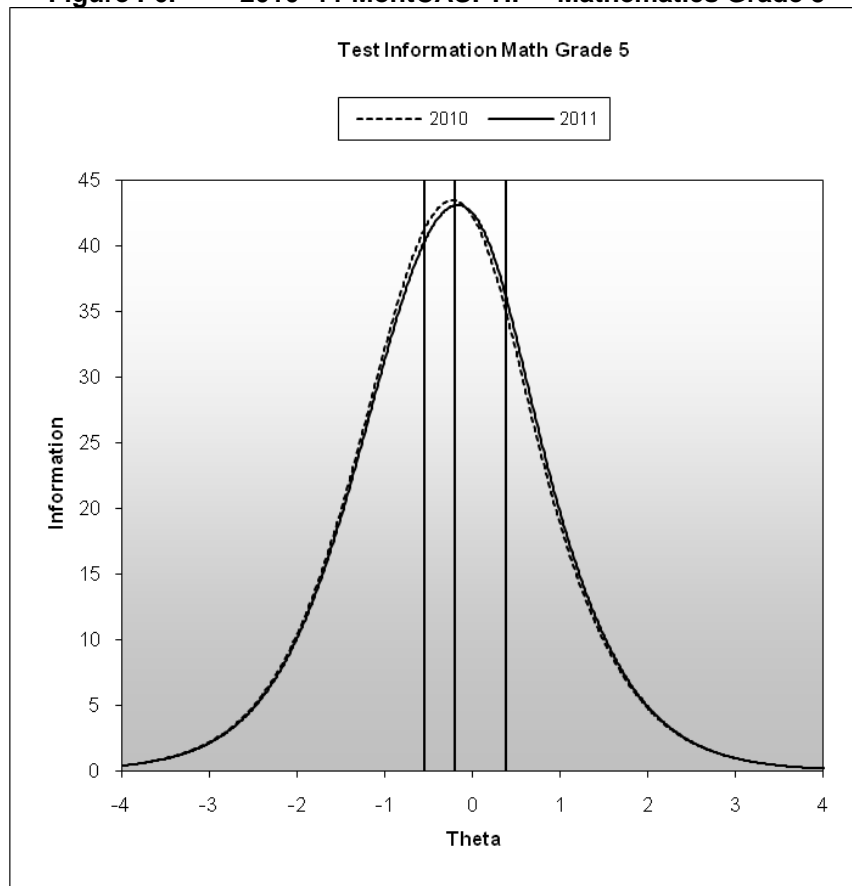


Figure I-7. 2010–11 MontCAS: TCC – Mathematics Grade 6

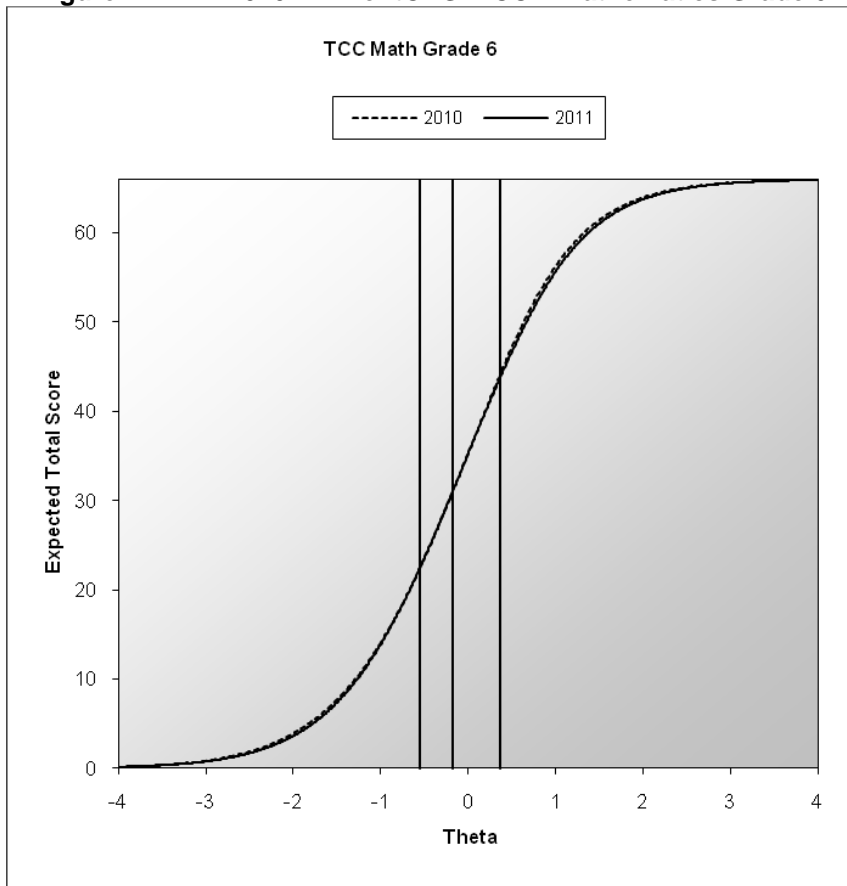


Figure I-8. 2010–11 MontCAS: TIF – Mathematics Grade 6

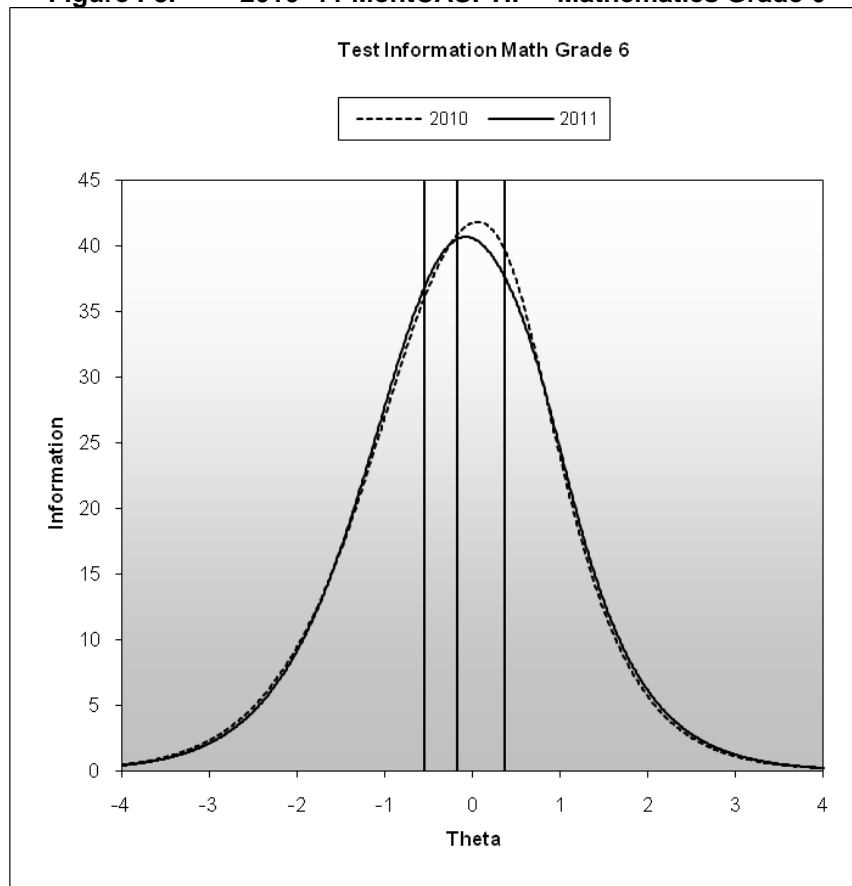


Figure I-9. 2010–11 MontCAS: TCC – Mathematics Grade 7

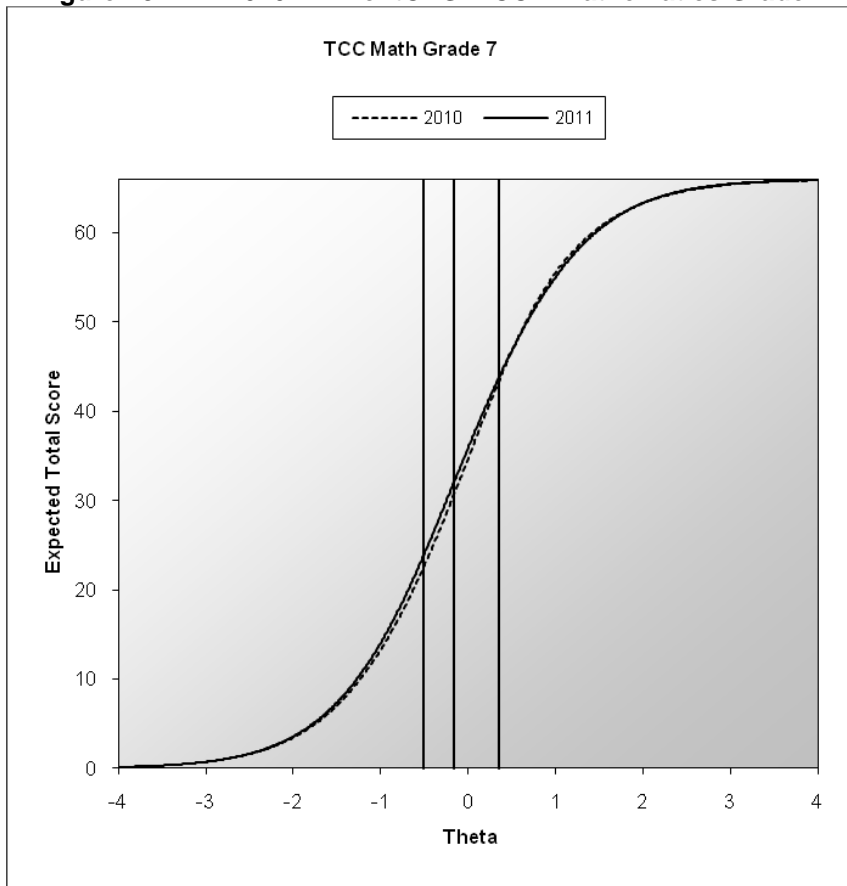


Figure I-10. 2010–11 MontCAS: TIF – Mathematics Grade 7

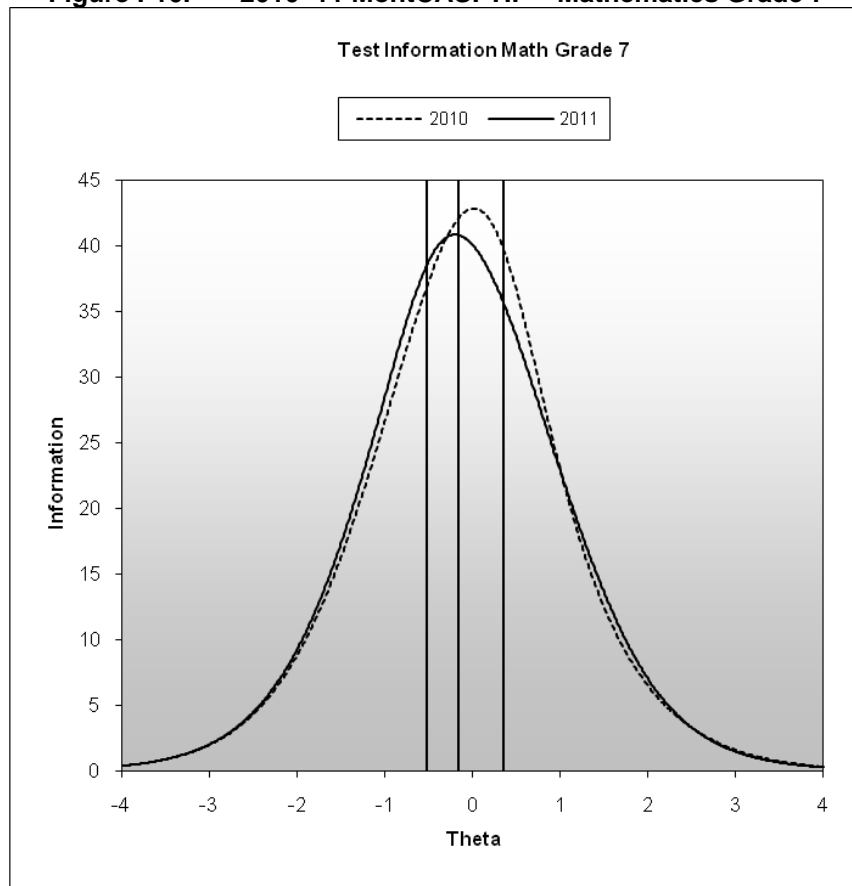


Figure I-11. 2010–11 MontCAS: TCC – Mathematics Grade 8

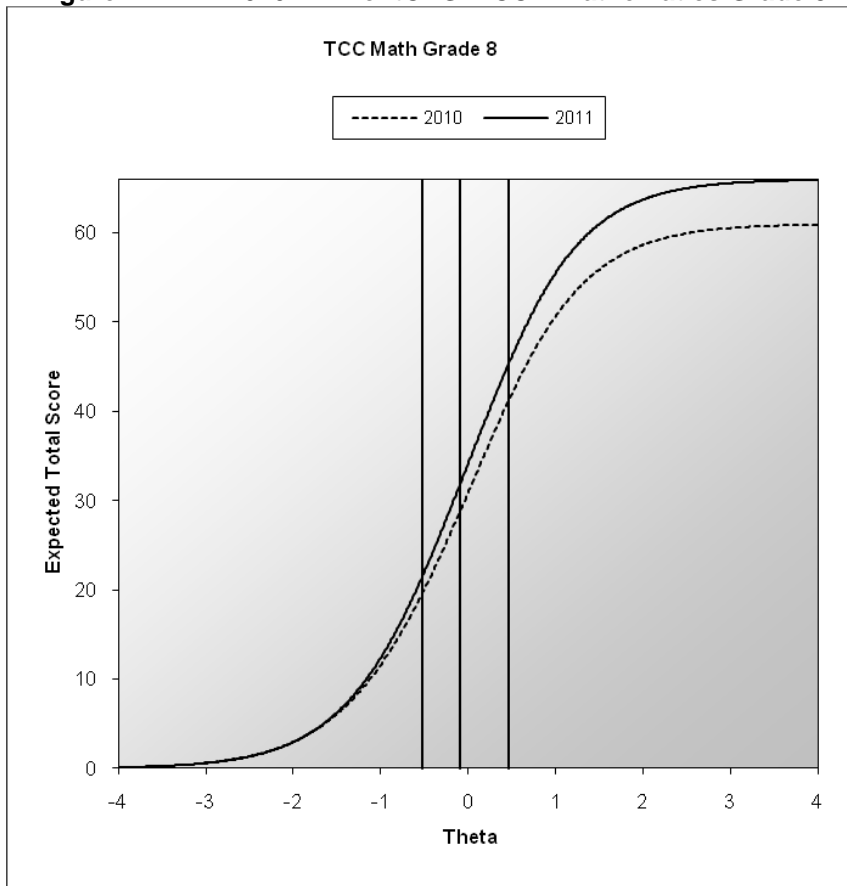


Figure I-12. 2010–11 MontCAS: TIF – Mathematics Grade 8

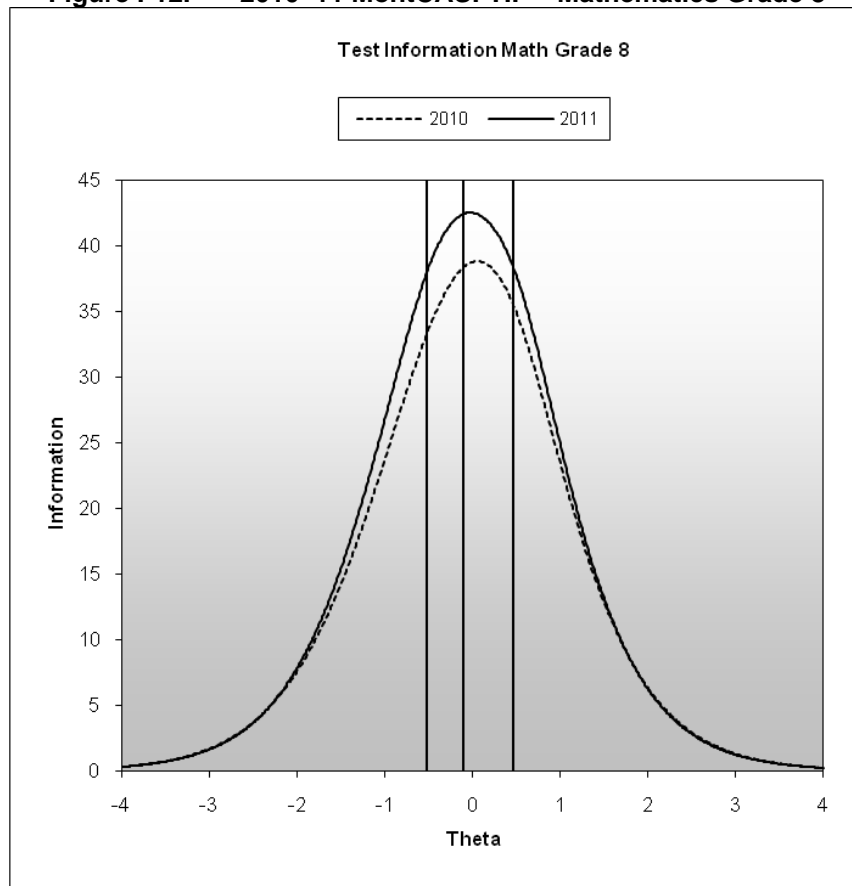


Figure I-13. 2010–11 MontCAS: TCC – Mathematics Grade 10

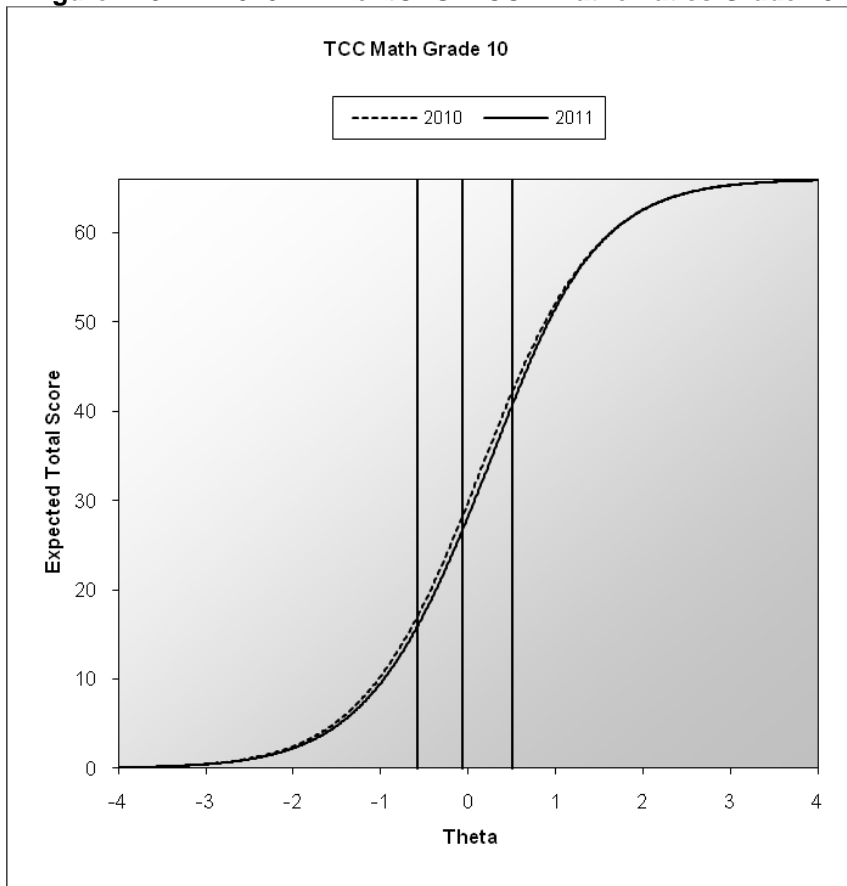


Figure I-14. 2010–11 MontCAS: TIF – Mathematics Grade 10

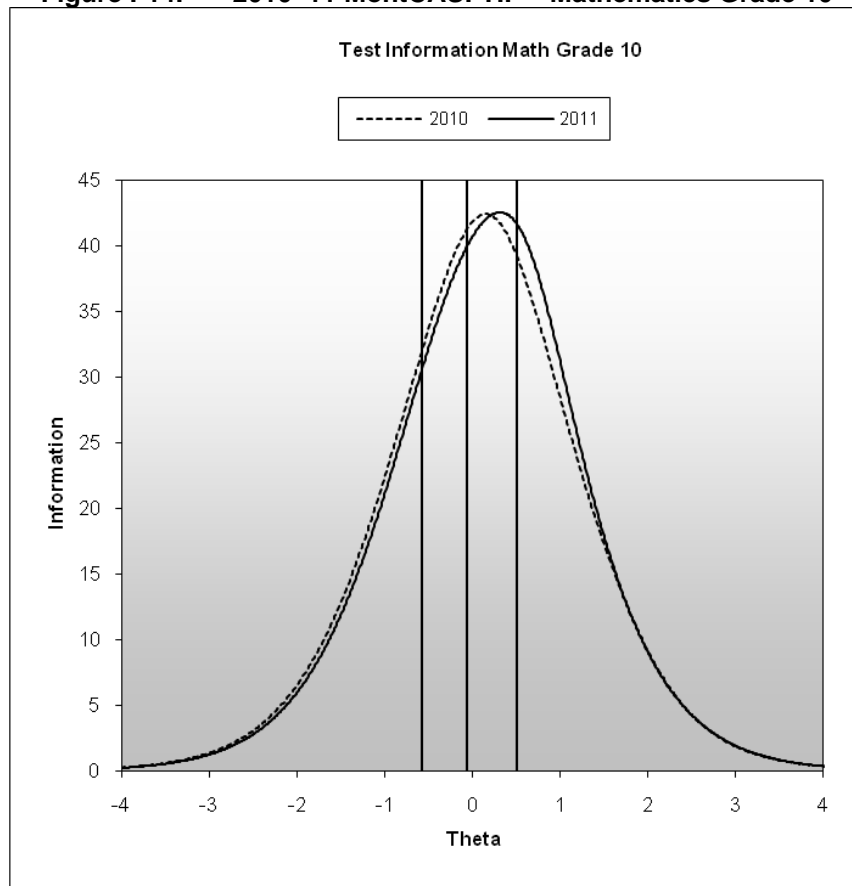


Figure I-15. 2010–11 MontCAS: TCC – Reading Grade 3

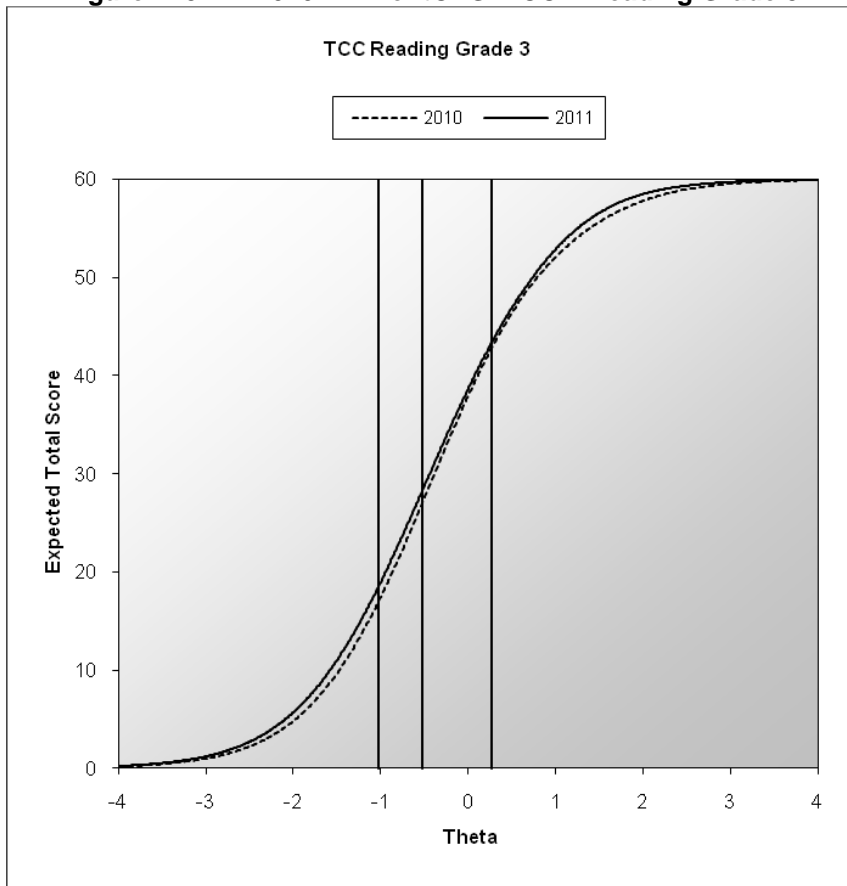


Figure I-16. 2010–11 MontCAS: TIF – Reading Grade 3

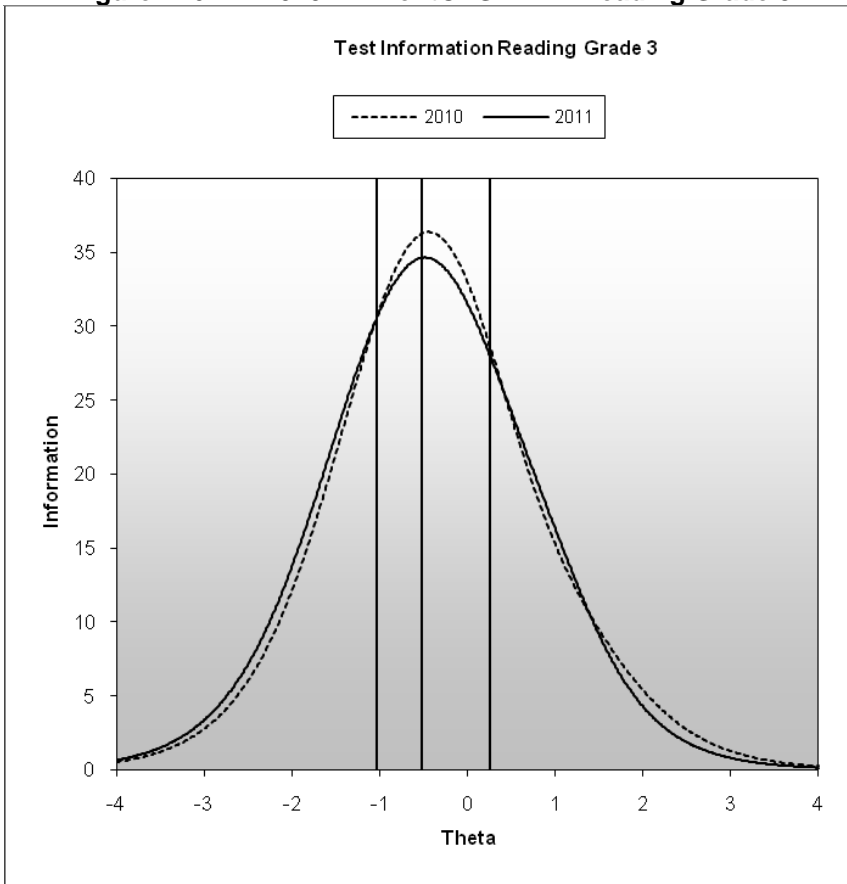


Figure I-17. 2010–11 MontCAS: TCC – Reading Grade 4

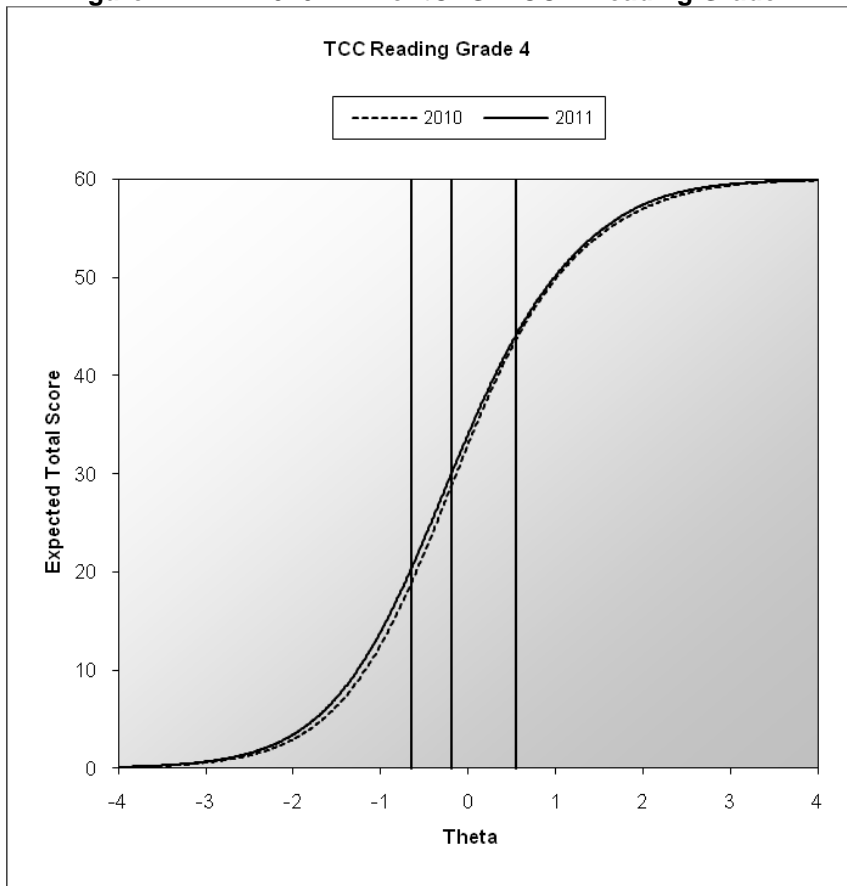


Figure I-18. 2010–11 MontCAS: TIF – Reading Grade 4

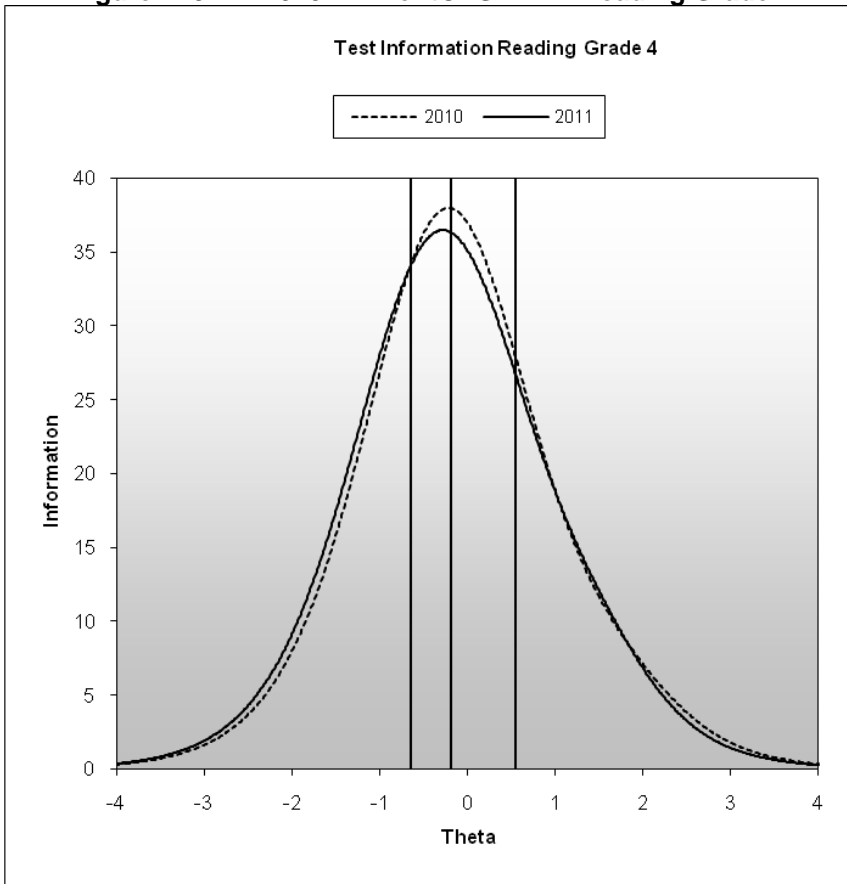


Figure I-19. 2010–11 MontCAS: TCC – Reading Grade 5

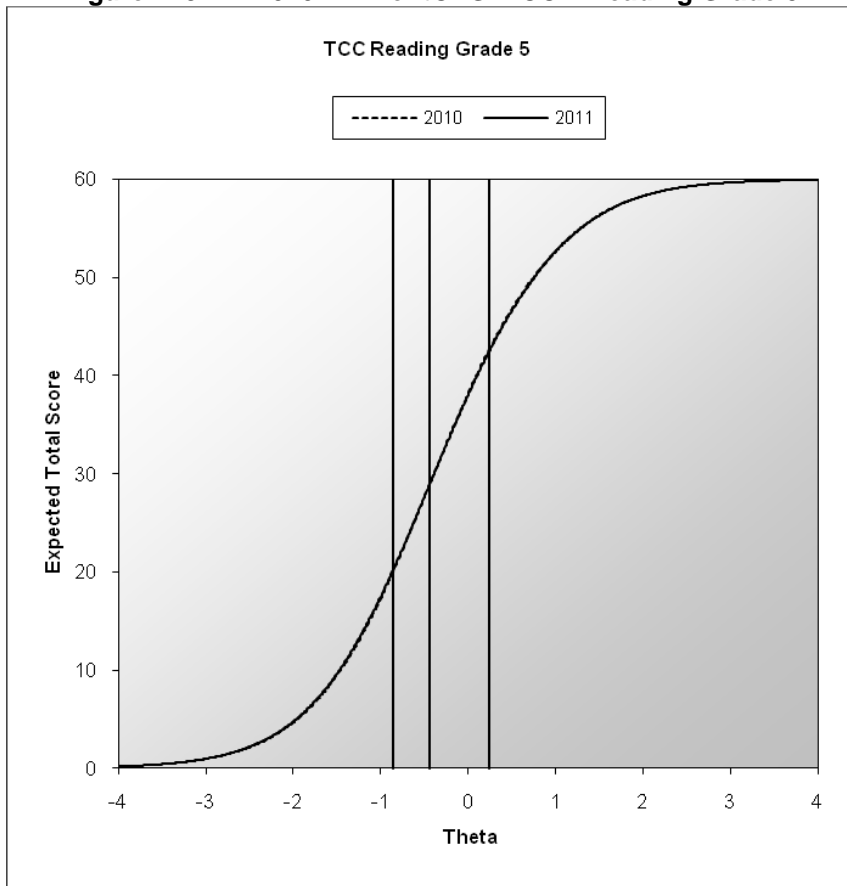


Figure I-20. 2010–11 MontCAS: TIF – Reading Grade 5

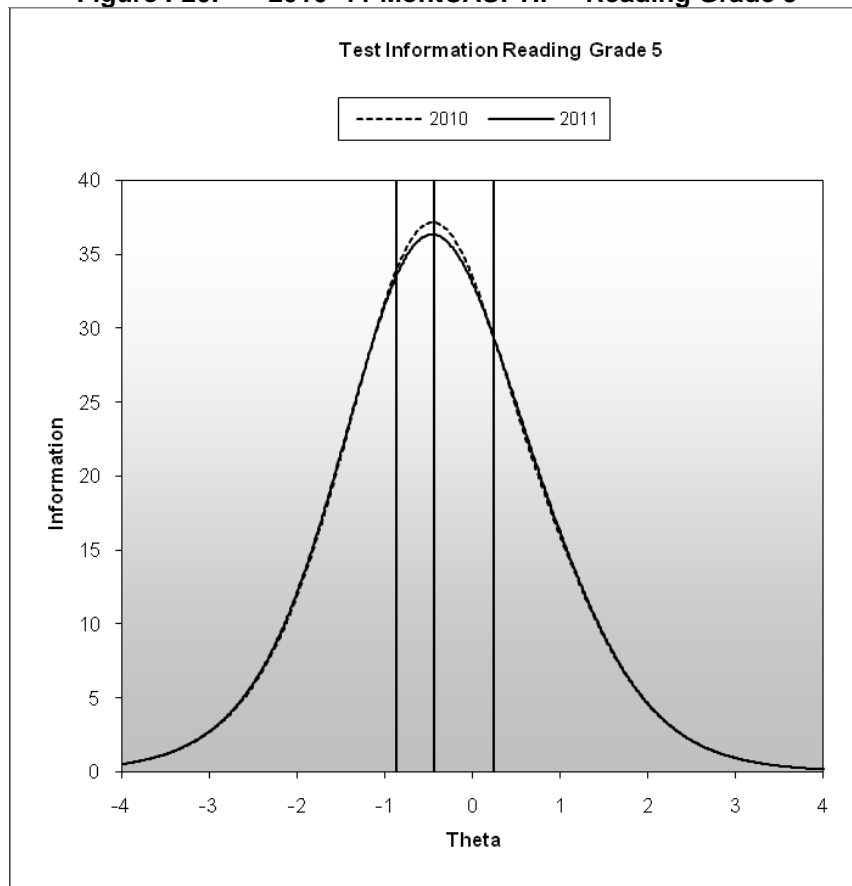


Figure I-21. 2010–11 MontCAS: TCC – Reading Grade 6

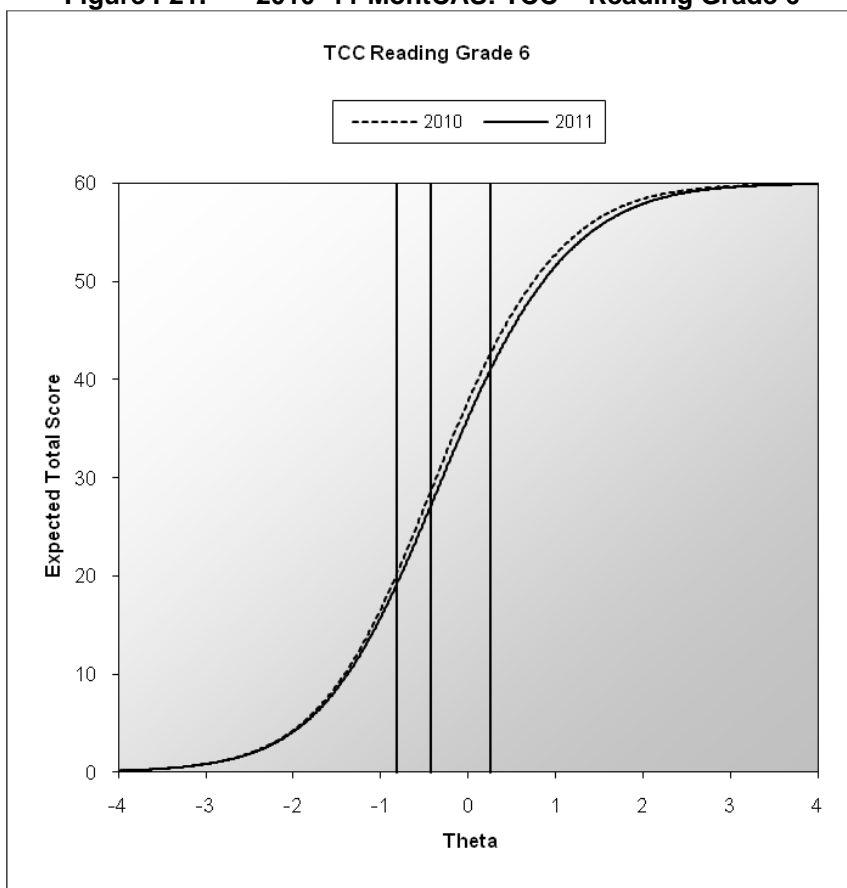


Figure I-22. 2010–11 MontCAS: TIF – Reading Grade 6

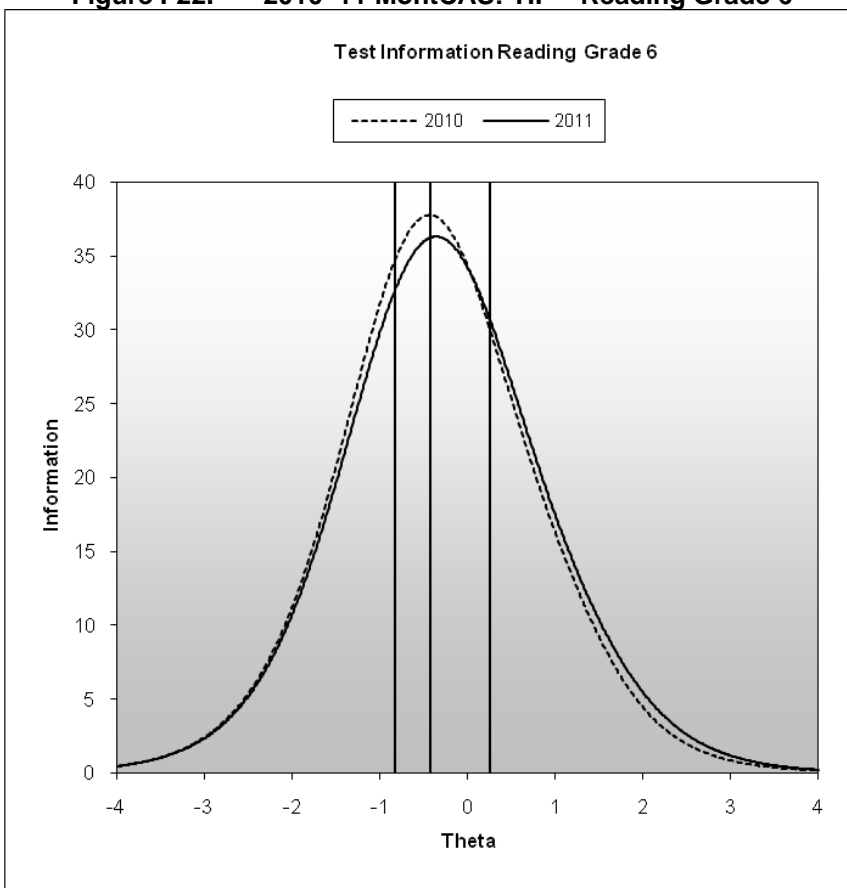


Figure I-23. 2010–11 MontCAS: TCC – Reading Grade 7

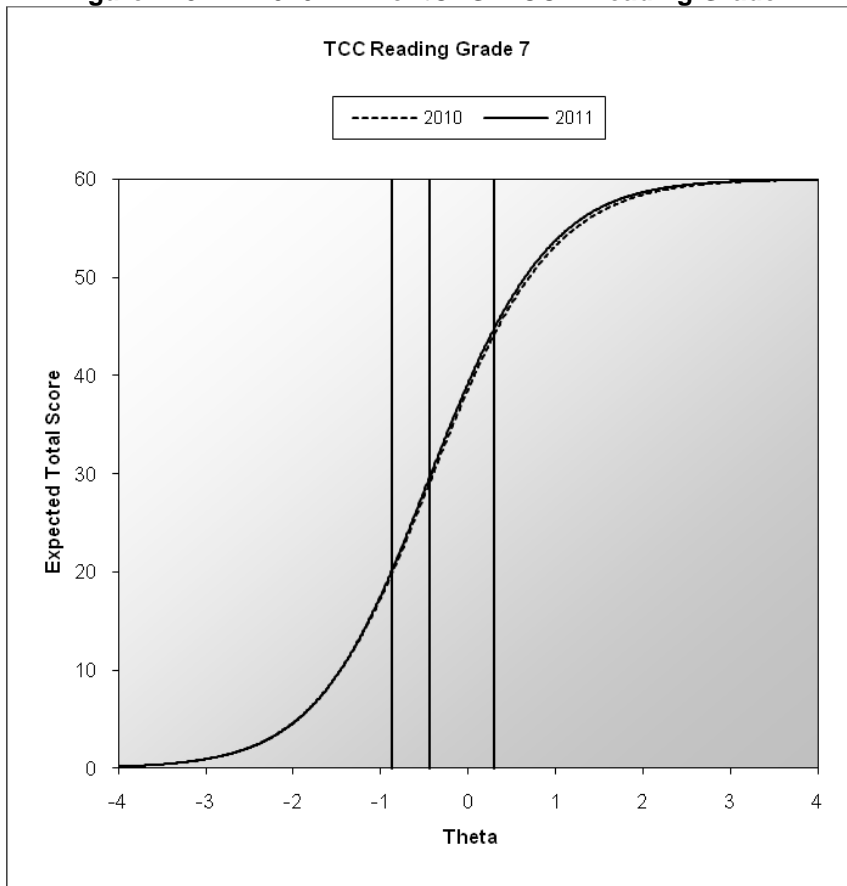


Figure I-24. 2010–11 MontCAS: TIF – Reading Grade 7

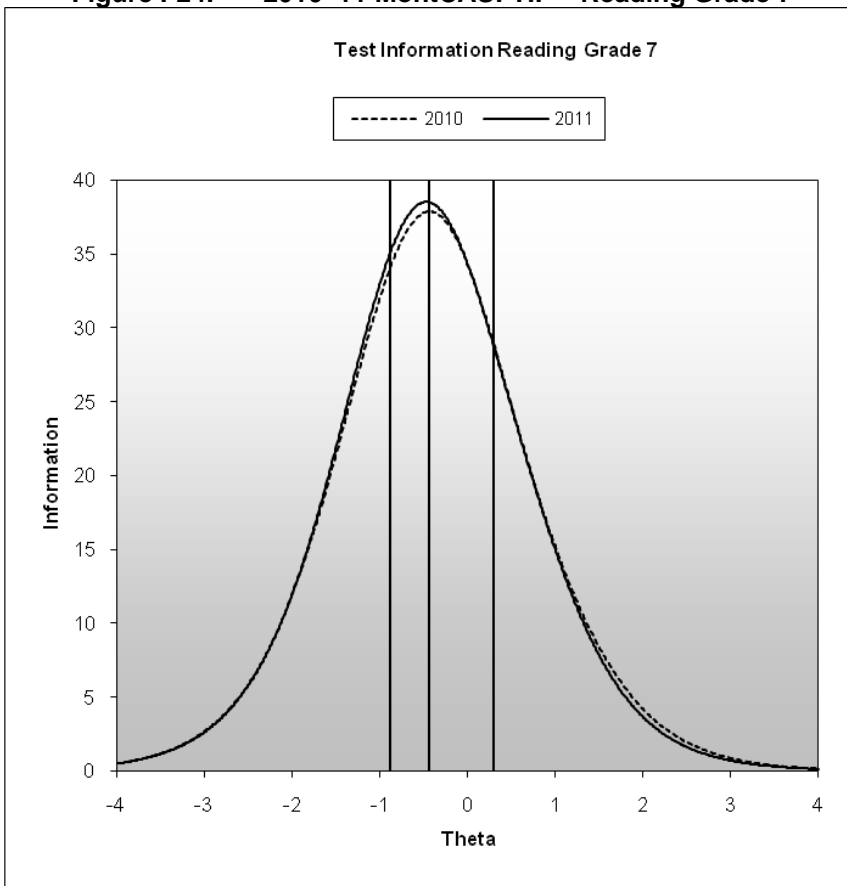


Figure I-25. 2010–11 MontCAS: TCC – Reading Grade 8

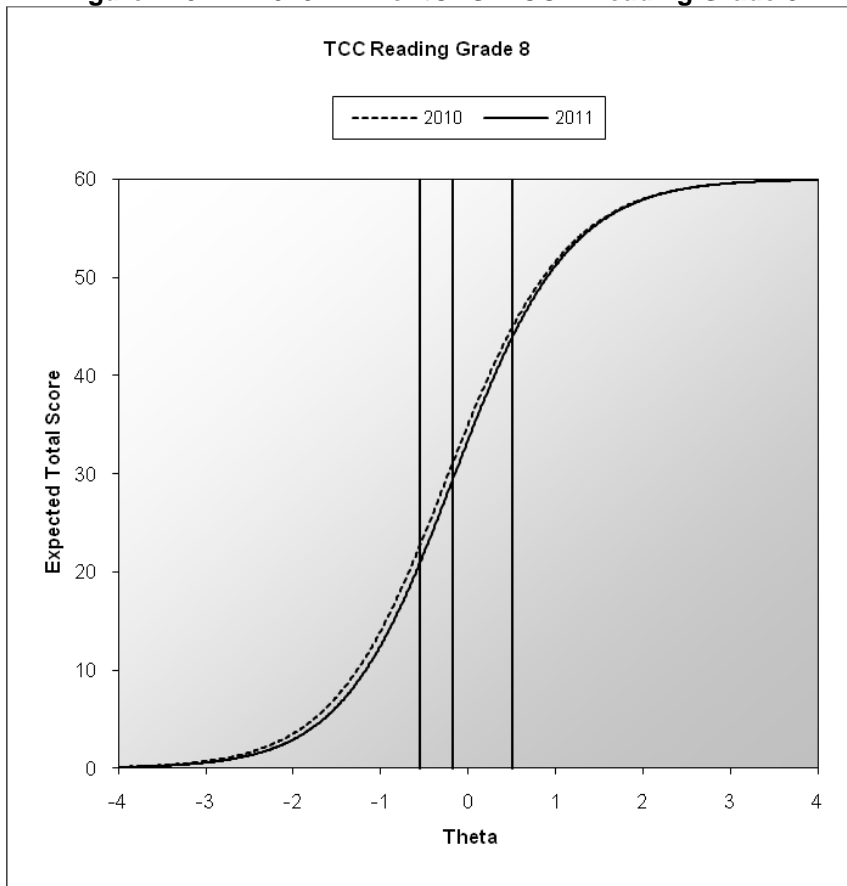


Figure I-26. 2010–11 MontCAS: TIF – Reading Grade 8

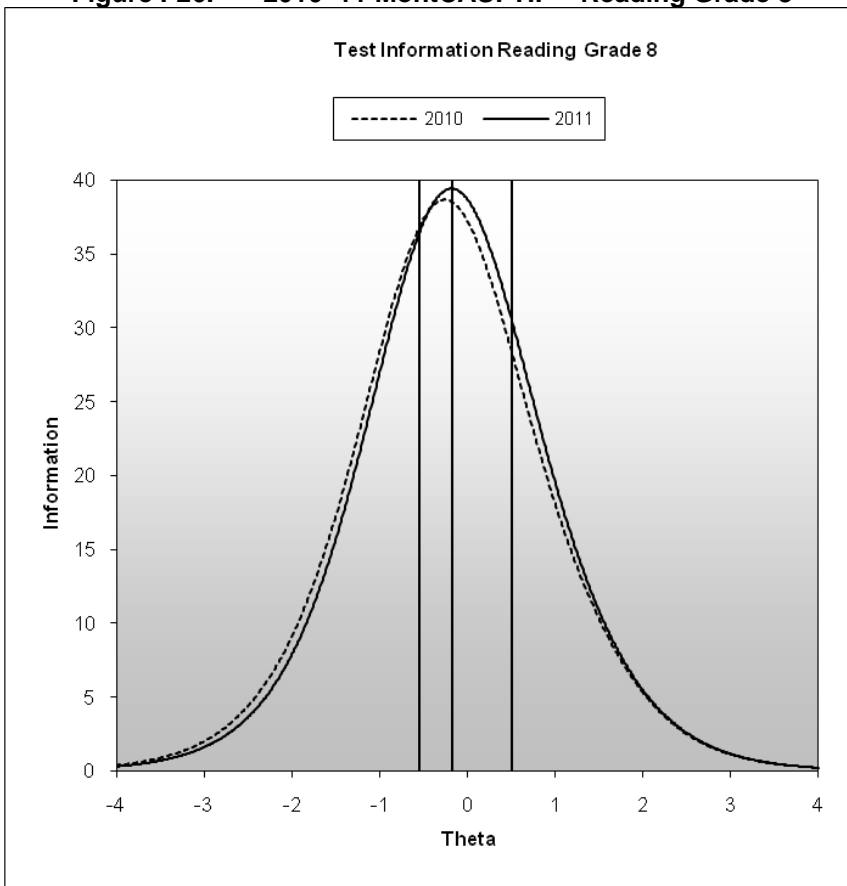


Figure I-27. 2010–11 MontCAS: TCC – Reading Grade 10

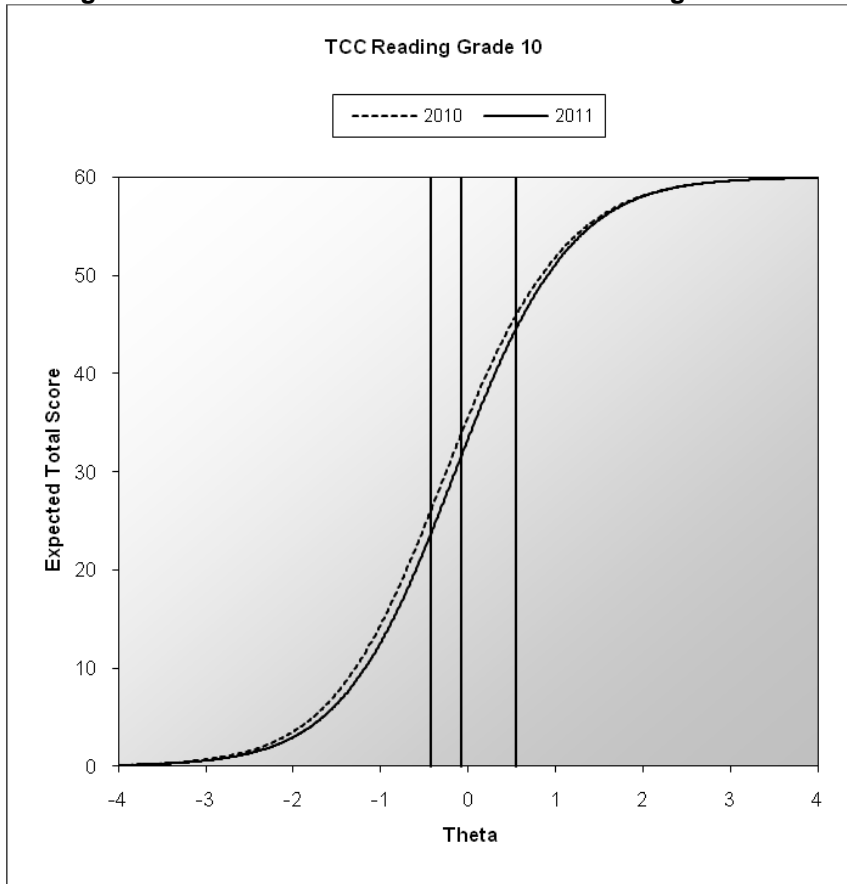


Figure I-28. 2010–11 MontCAS: TIF – Reading Grade 10

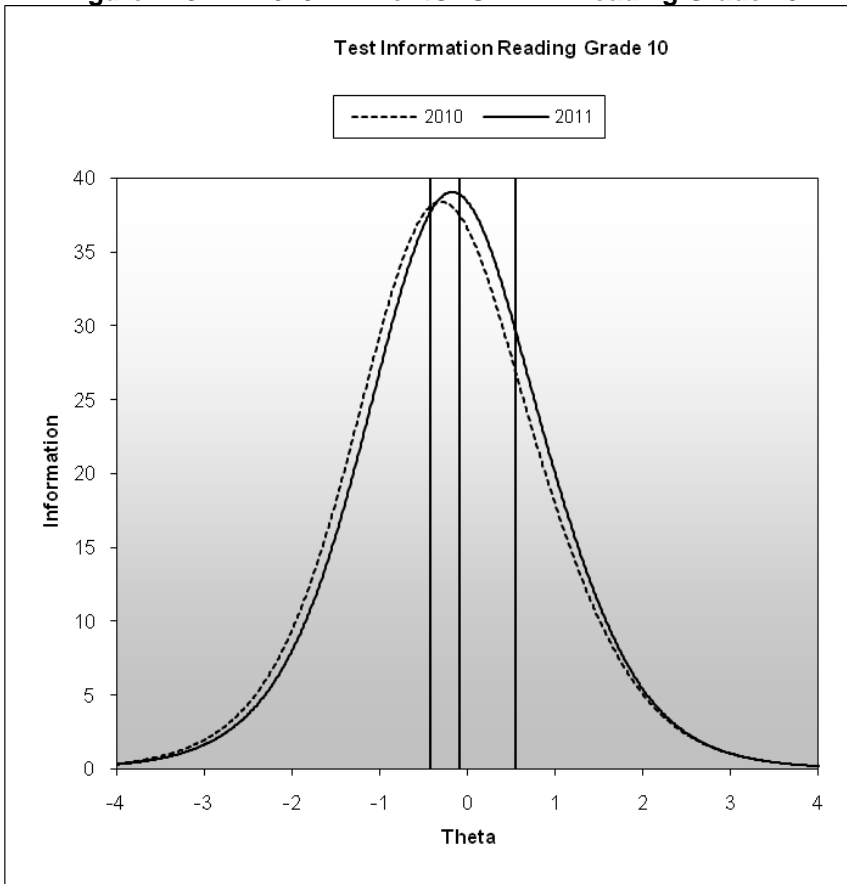


Figure I-29. 2010–11 MontCAS: TCC – Science Grade 4

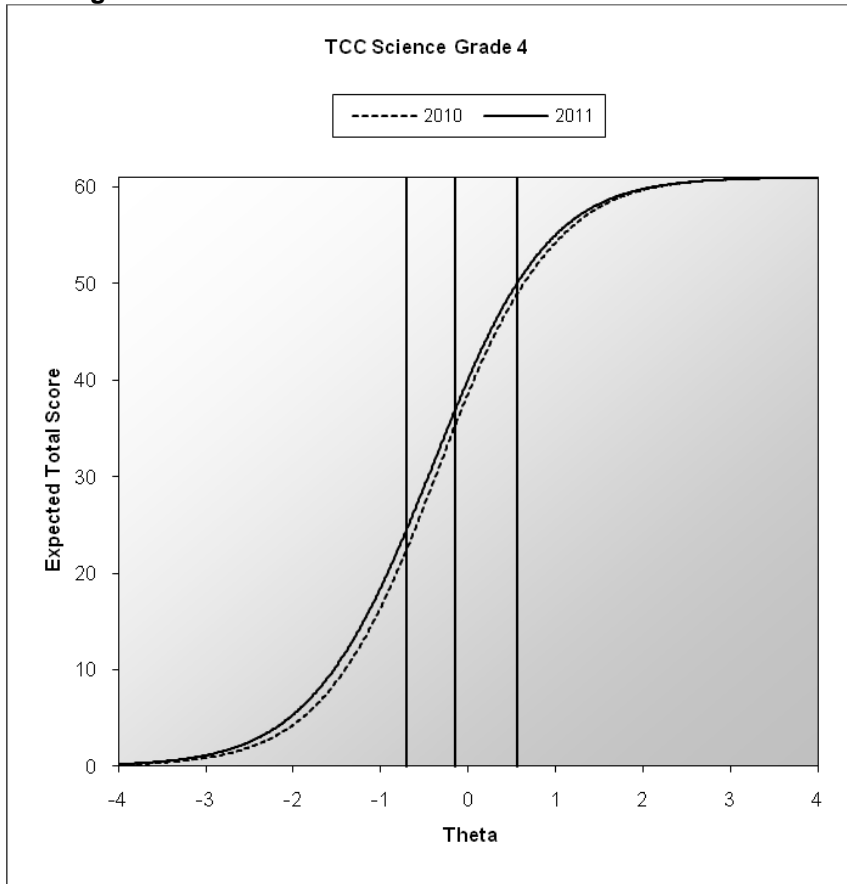


Figure I-30. 2010–11 MontCAS: TIF – Science Grade 4

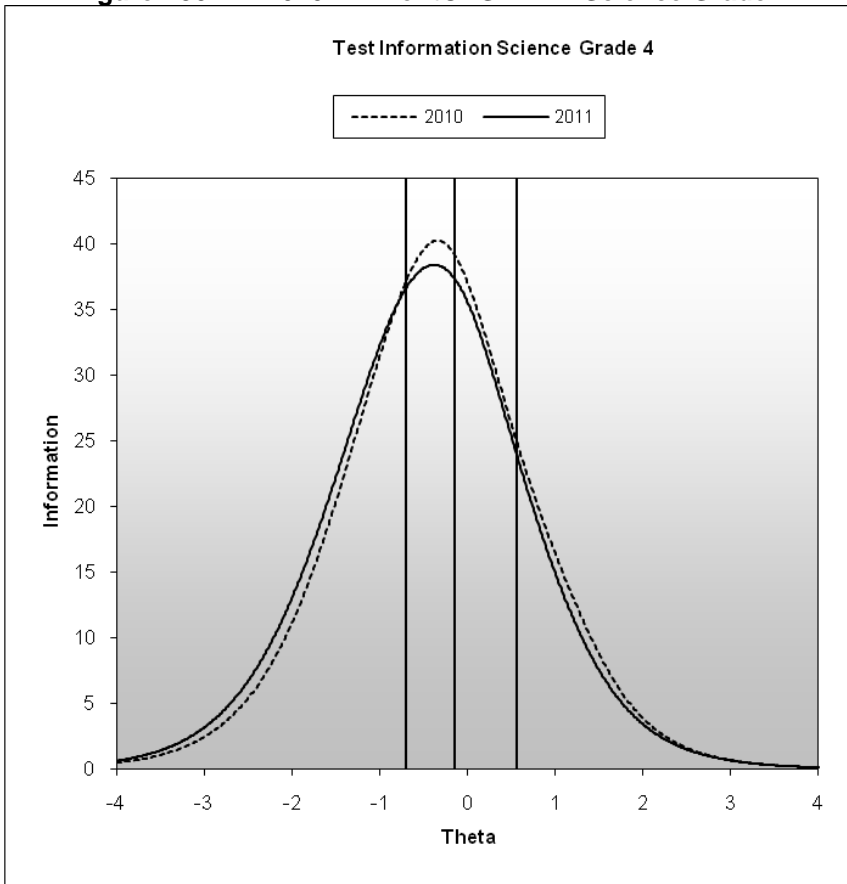


Figure I-31. 2010–11 MontCAS: TCC – Science Grade 8

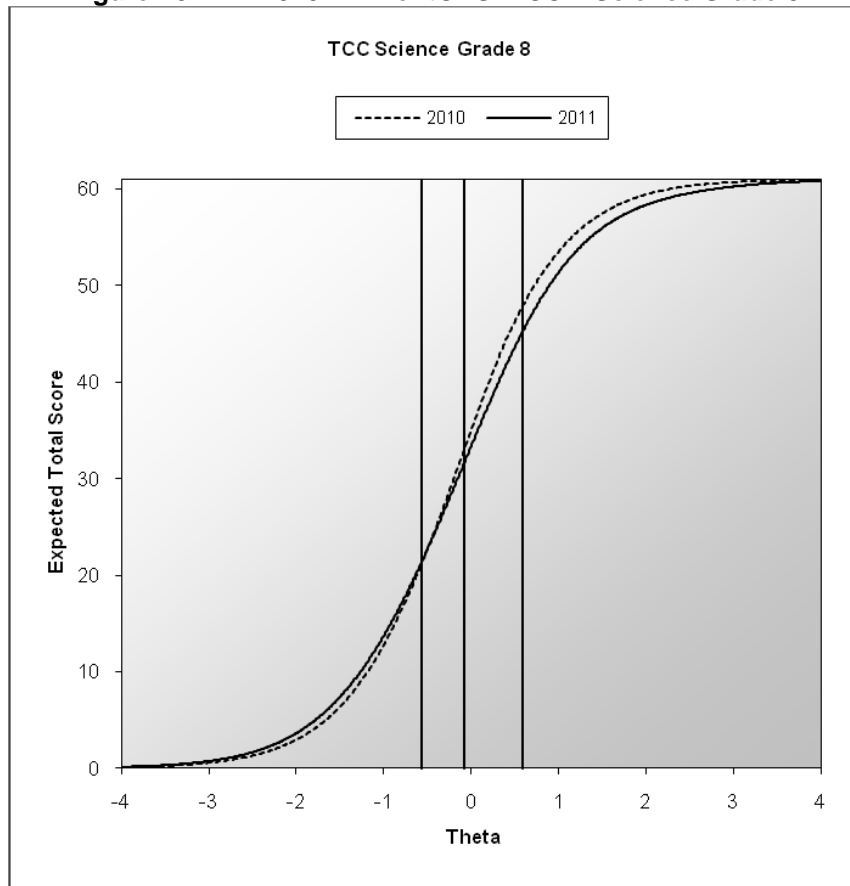


Figure I-32. 2010–11 MontCAS: TIF – Science Grade 8

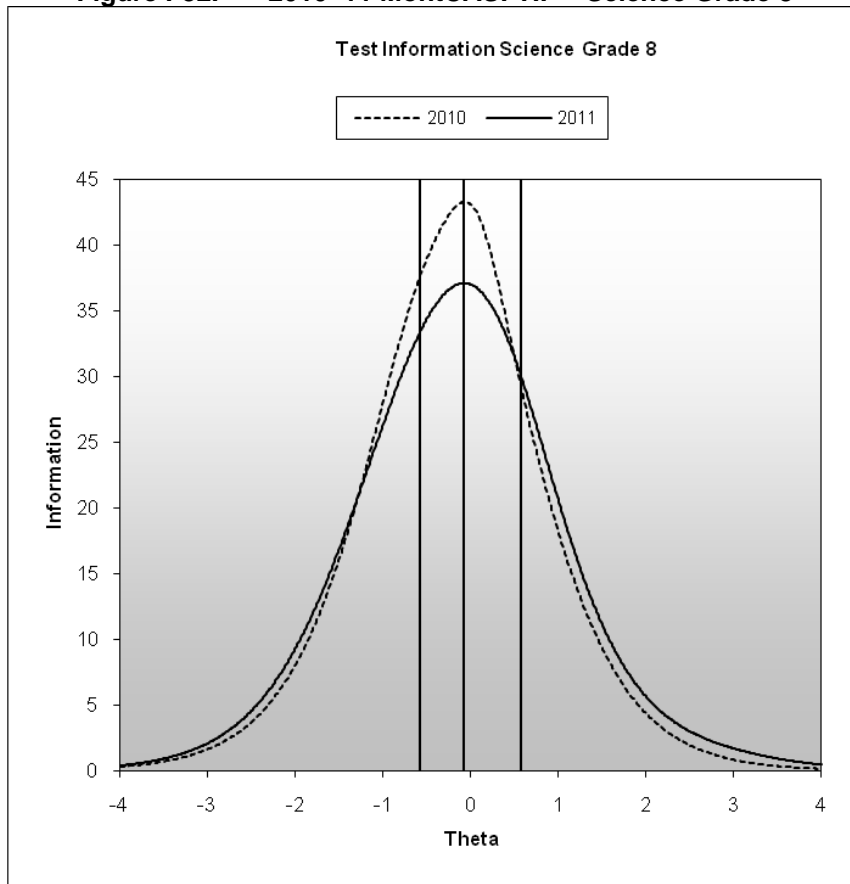


Figure I-33. 2010–11 MontCAS: TCC – Science Grade 10

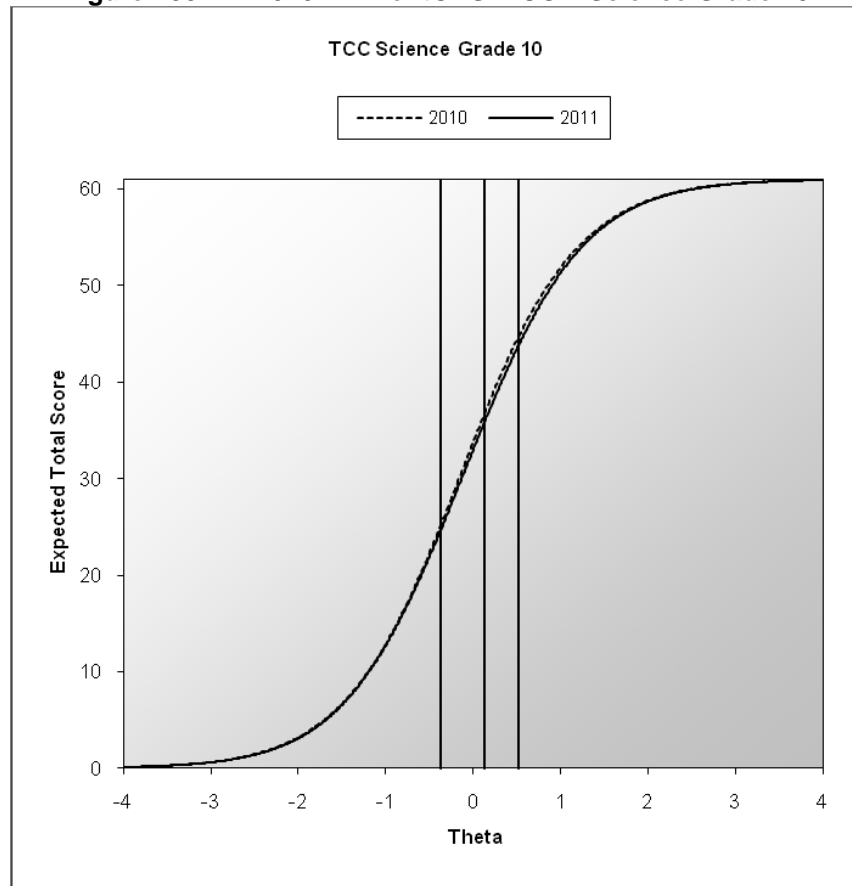
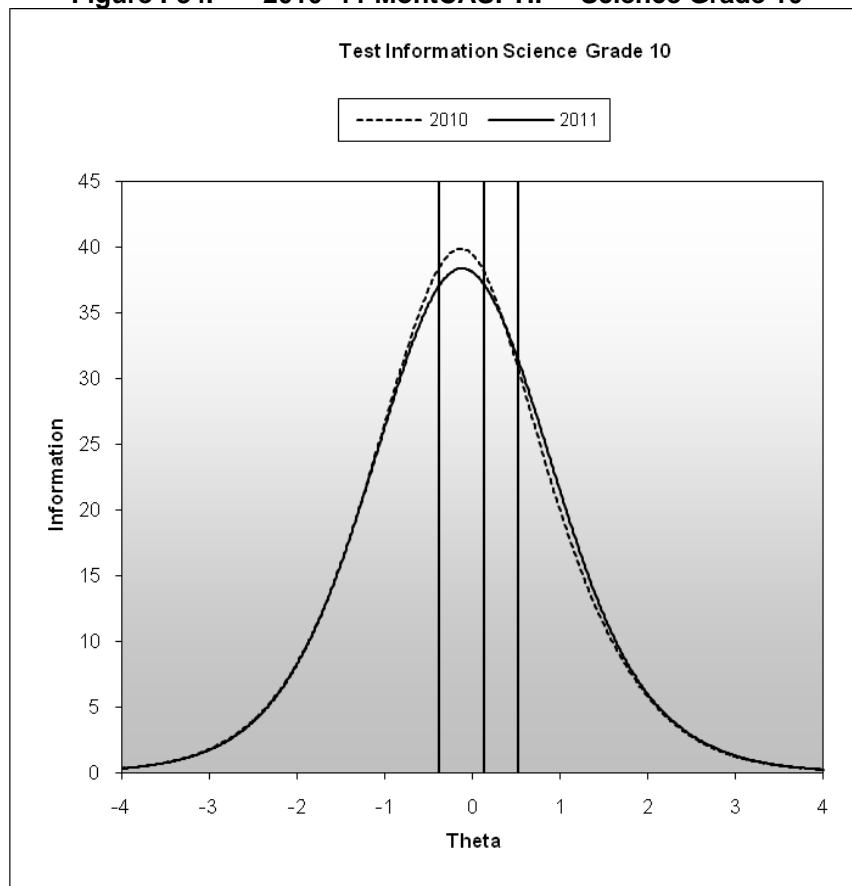


Figure I-34. 2010–11 MontCAS: TIF – Science Grade 10



Appendix J—*B*-PLOTS

Figure J-1. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 3

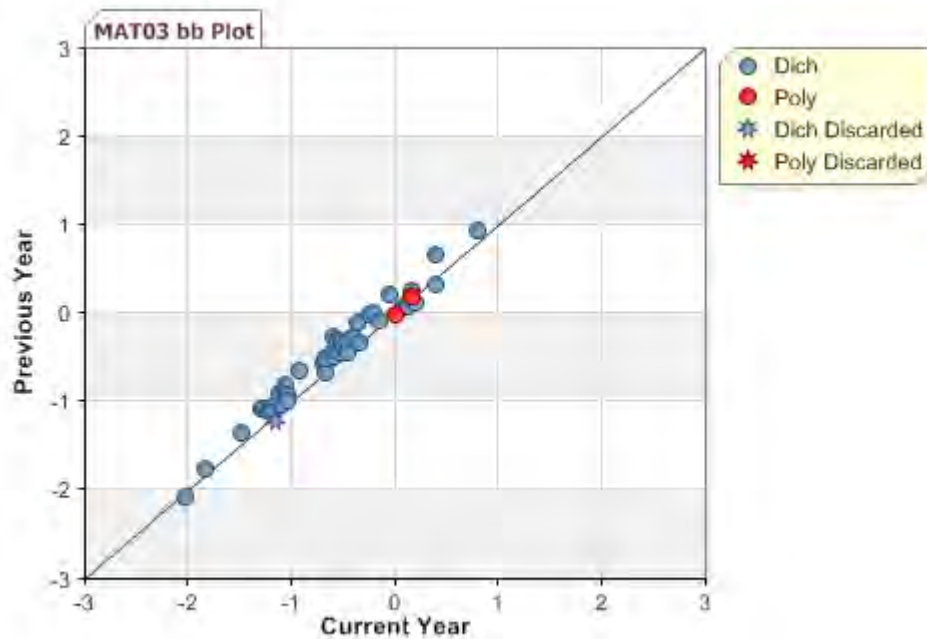


Figure J-2. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 4

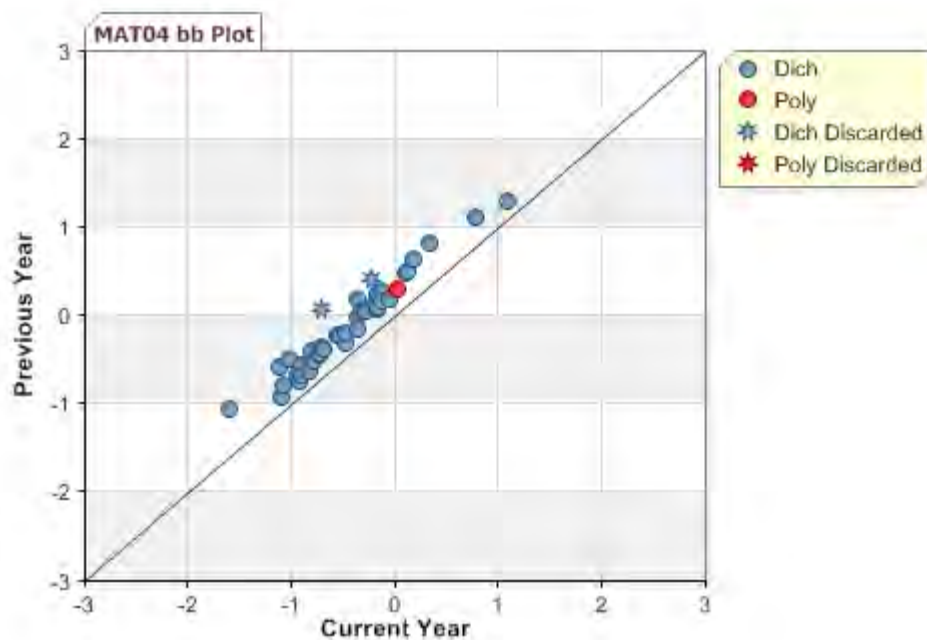


Figure J-3. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 5

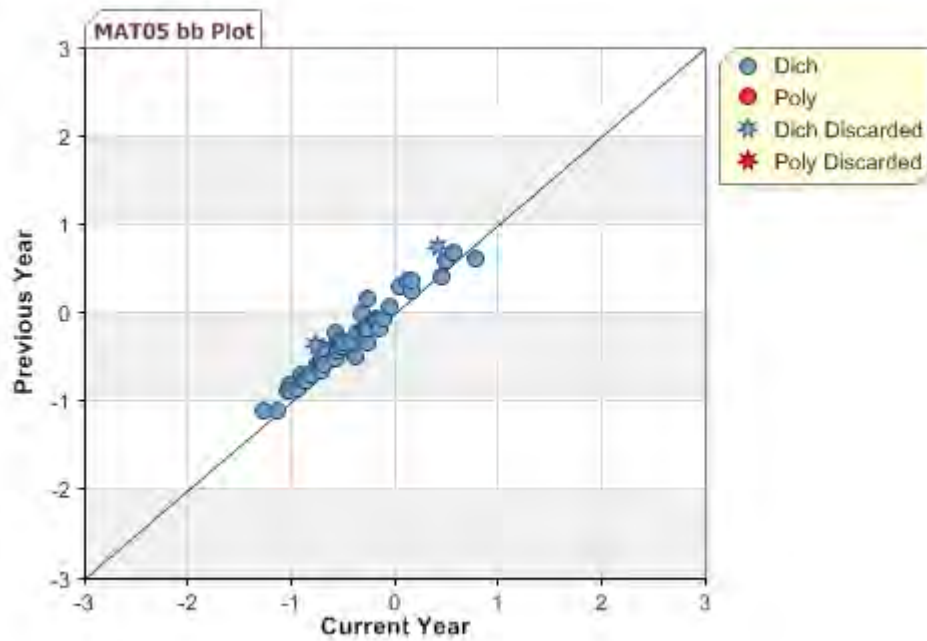


Figure J-4. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 6

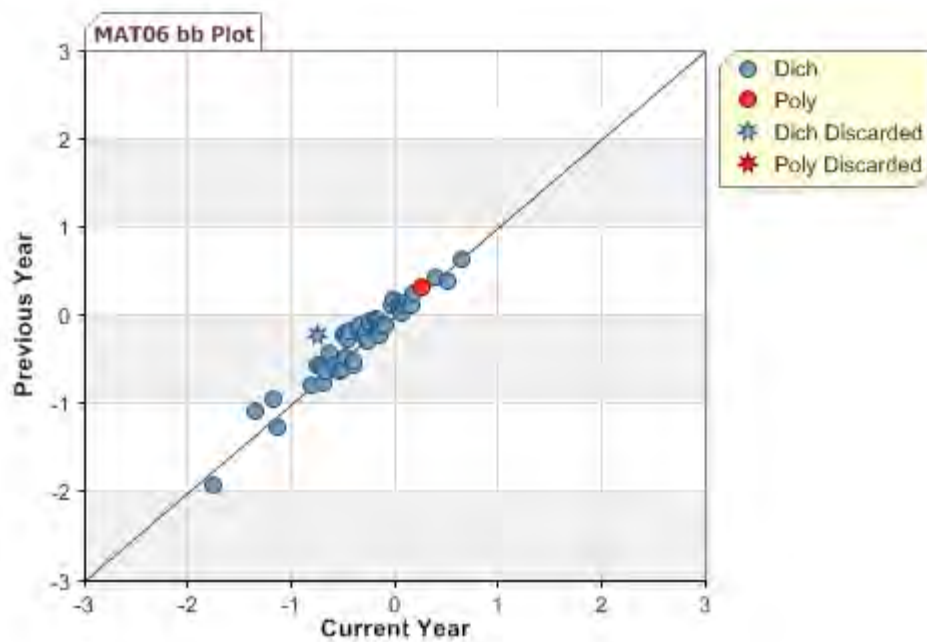


Figure J-5. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 7

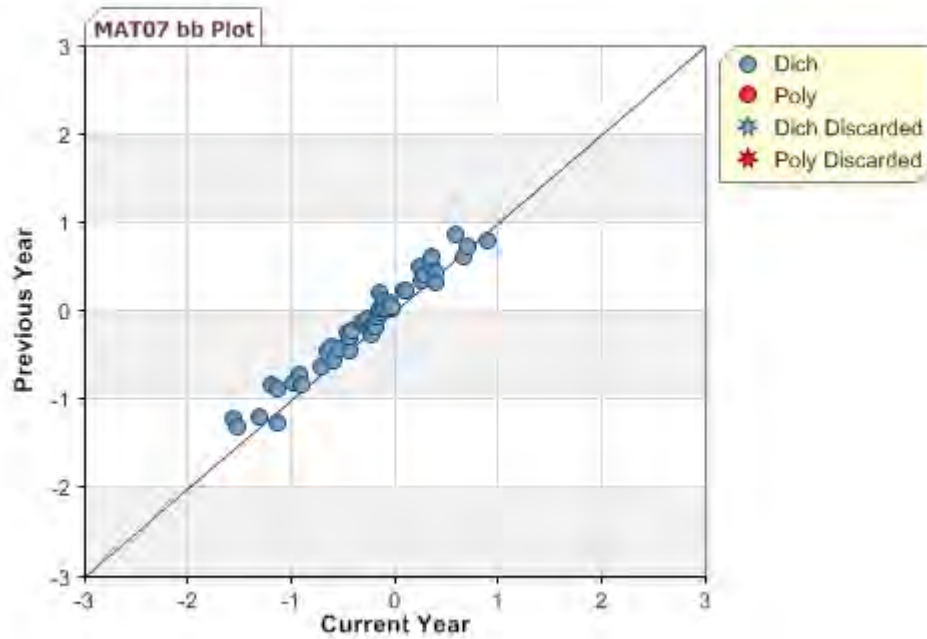


Figure J-6. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 8

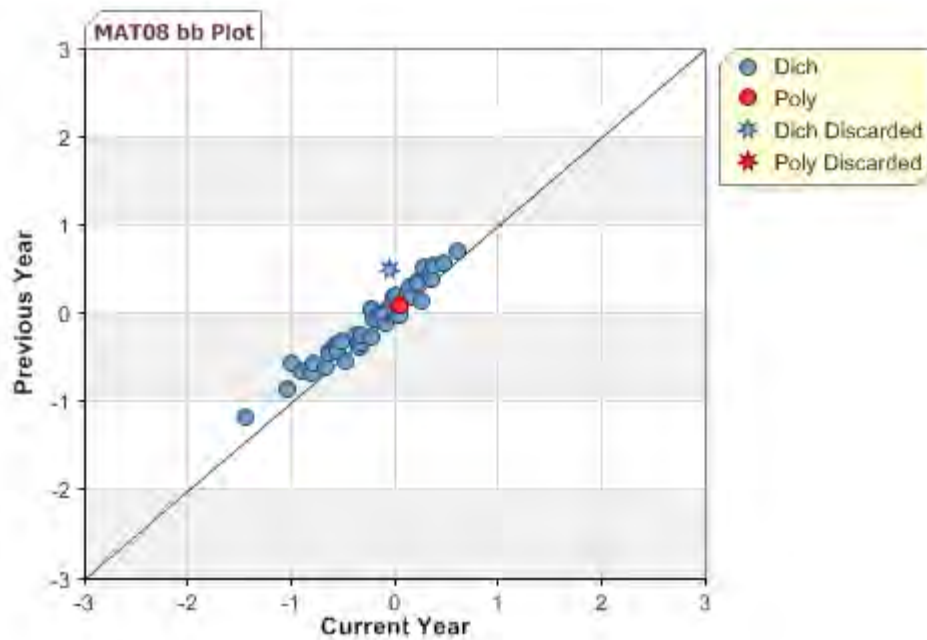


Figure J-7. 2010–11 MontCAS: *b*-Plots – Mathematics Grade 10

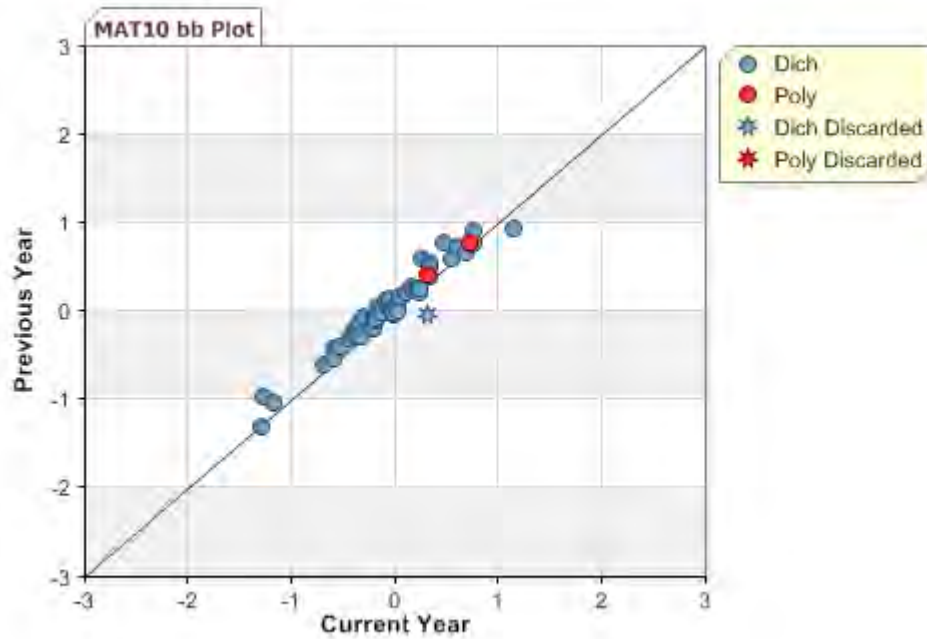


Figure J-8. 2010–11 MontCAS: *b*-Plots – Reading Grade 3

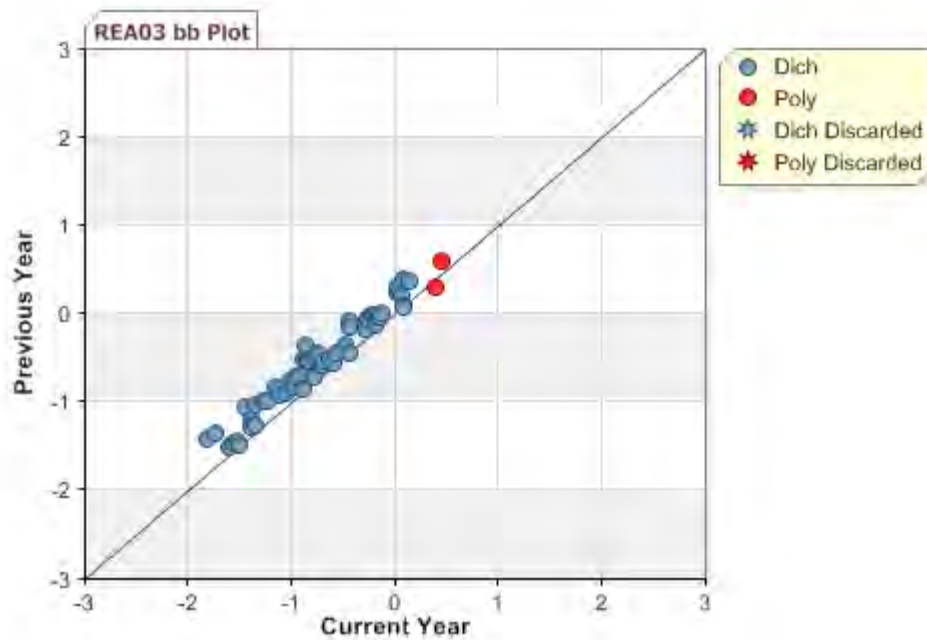


Figure J-9. 2010–11 MontCAS: *b*-Plots – Reading Grade 4

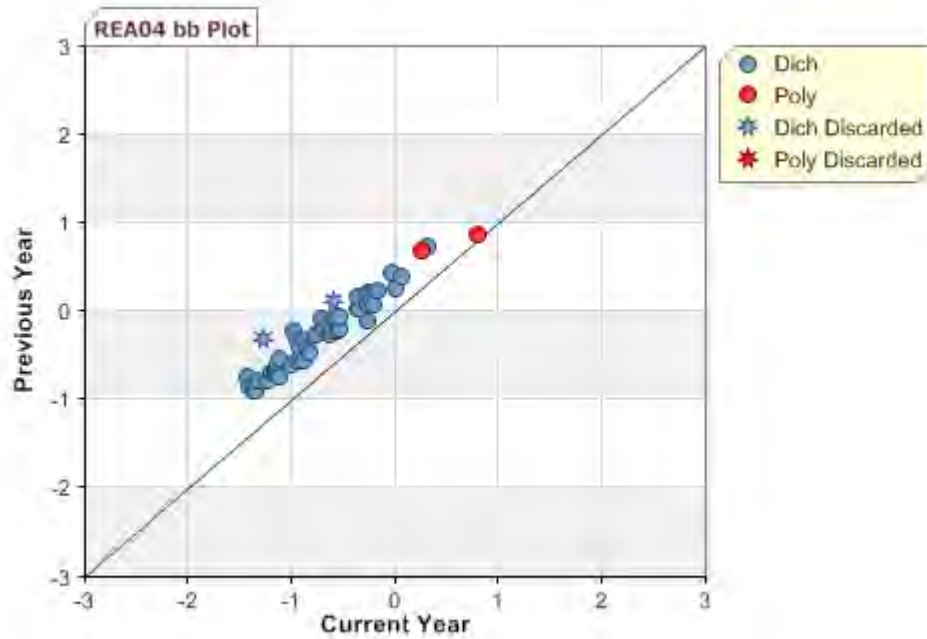


Figure J-10. 2010–11 MontCAS: *b*-Plots – Reading Grade 5

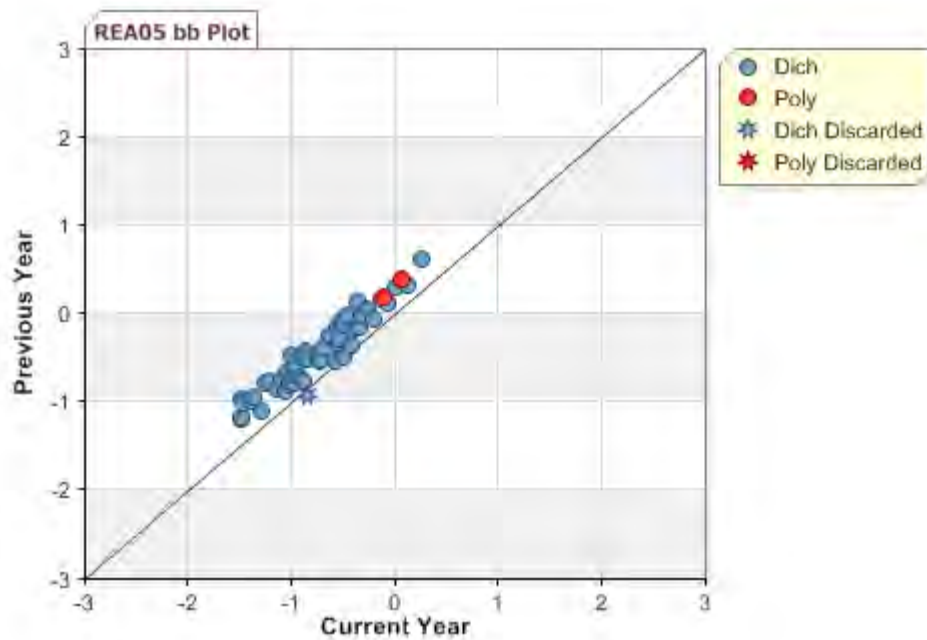


Figure J-11. 2010–11 MontCAS: *b*-Plots – Reading Grade 6

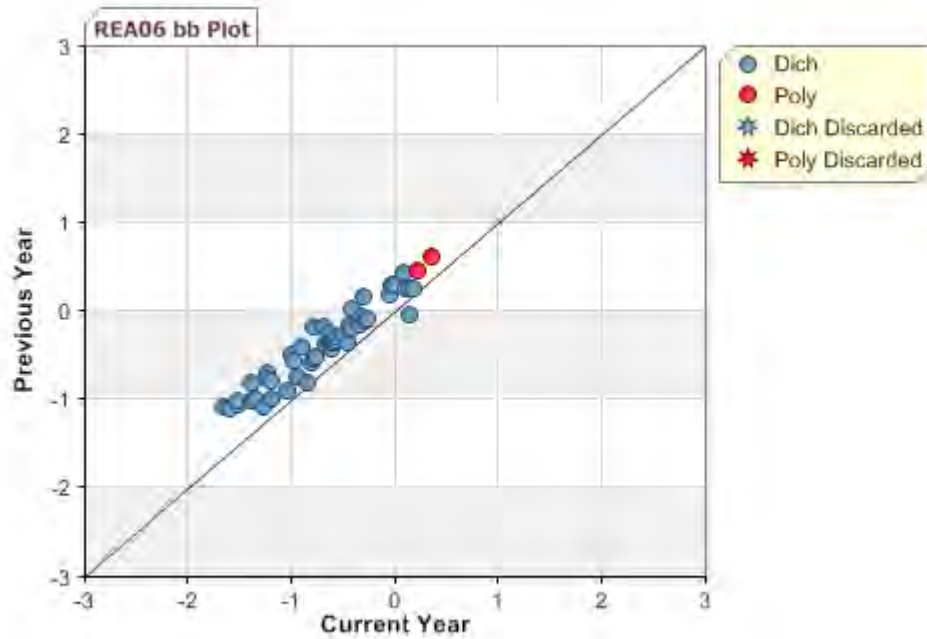


Figure J-12. 2010–11 MontCAS: *b*-Plots – Reading Grade 7

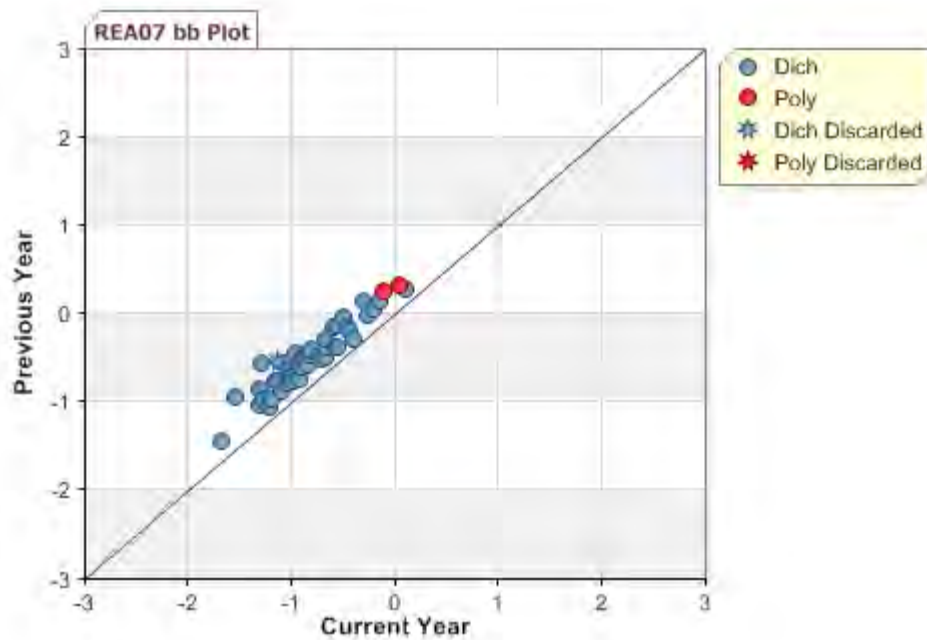


Figure J-13. 2010–11 MontCAS: *b*-Plots – Reading Grade 8

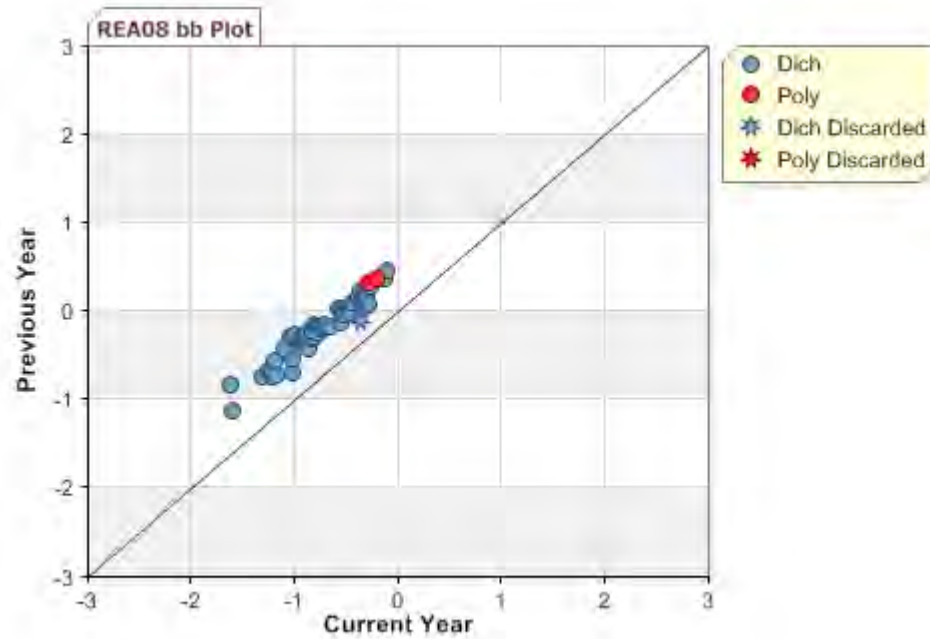


Figure J-14. 2010–11 MontCAS: *b*-Plots – Reading Grade 10

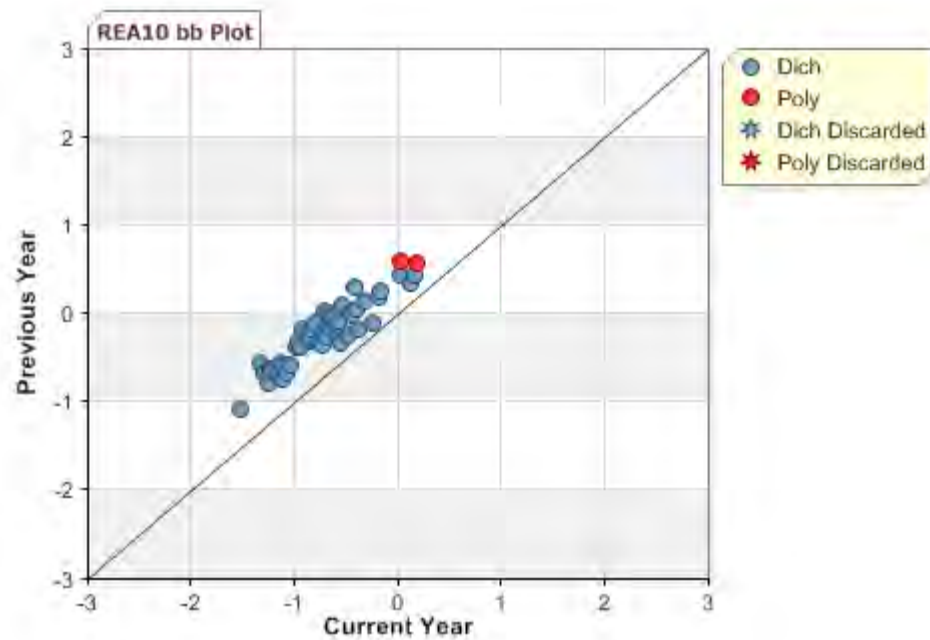


Figure J-15. 2010–11 MontCAS: *b*-Plots – Science Grade 4

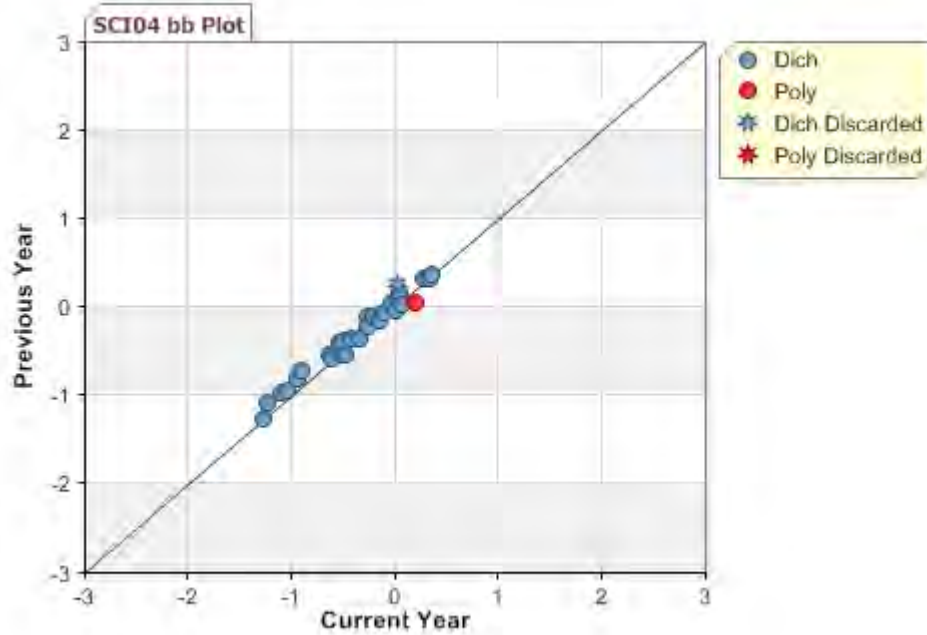


Figure J-16. 2010–11 MontCAS: *b*-Plots – Science Grade 8

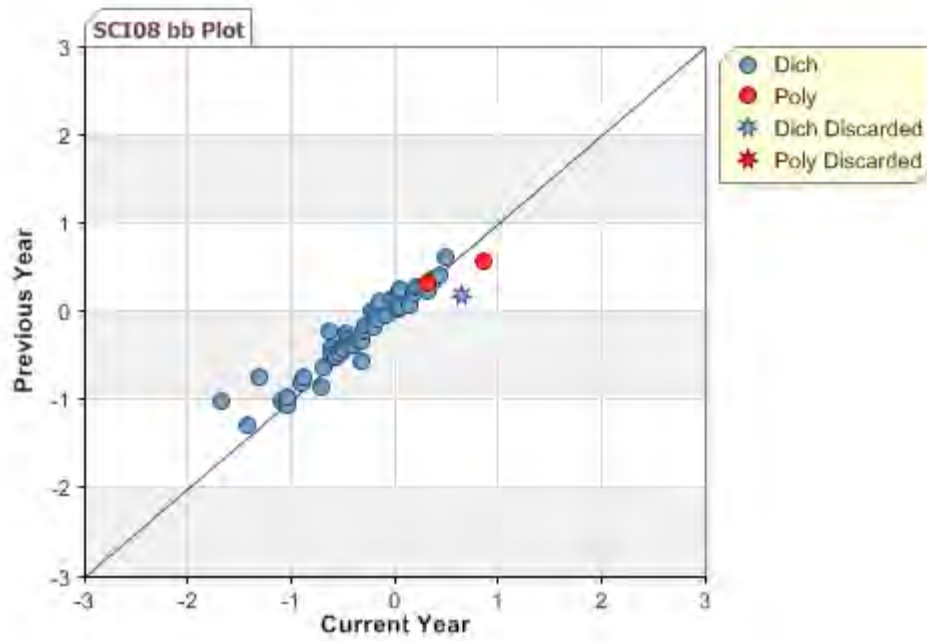
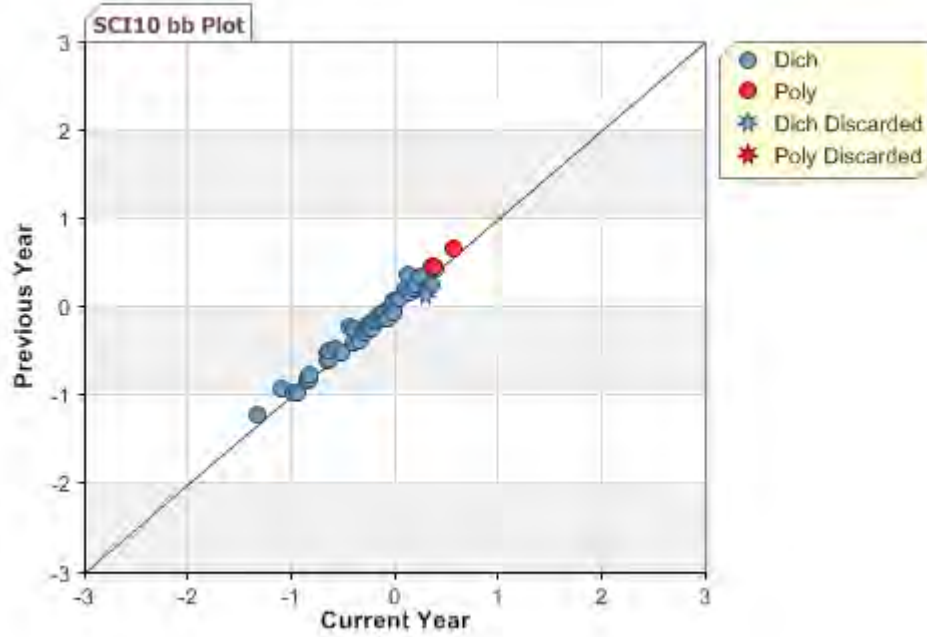


Figure J-17. 2010–11 MontCAS: *b*-Plots – Science Grade 10



Appendix K—ANALYSES OF EQUATING ITEMS (DELTA & RESCORE ANALYSES)

Table K-1. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 3

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
138751	0.80	0.82	9.63	9.34	1	False	-0.32
138755	0.65	0.69	11.46	11.02	1	False	0.21
138763	0.84	0.82	9.02	9.34	1	False	0.54
138765	0.63	0.64	11.67	11.57	1	False	-1.25
138780	0.66	0.68	11.35	11.13	1	False	-0.87
138781	0.80	0.83	9.63	9.18	1	False	0.44
138787	0.64	0.69	11.57	11.02	1	False	0.72
138789	0.52	0.49	12.80	13.10	1	False	0.86
138795	0.70	0.66	10.90	11.35	1	False	1.38
138799	0.46	0.44	13.40	13.60	4	False	0.44
138799	0.44	0.44	13.60	13.60	4	False	-0.53
138832	0.71	0.73	10.79	10.55	1	False	-0.72
138865	0.32	0.38	14.87	14.22	1	False	0.86
138879	0.73	0.73	10.55	10.55	1	False	-0.85
138917	0.82	0.82	9.34	9.34	1	False	-0.97
138929	0.80	0.82	9.63	9.34	1	False	-0.32
138982	0.84	0.87	9.02	8.49	1	False	0.88
138986	0.55	0.59	12.50	12.09	1	False	-0.07
138995	0.64	0.70	11.57	10.90	1	False	1.28
139015	0.69	0.70	11.02	10.90	1	False	-1.35
139031	0.57	0.56	12.29	12.40	1	False	-0.17
139045	0.86	0.86	8.68	8.68	1	False	-1.04
139051	0.52	0.58	12.80	12.19	1	False	0.87
242748	0.89	0.84	8.09	9.02	1	True	3.44
34618	0.77	0.79	10.04	9.77	1	False	-0.49
42981	0.88	0.90	8.30	7.87	1	False	0.46
43120	0.93	0.93	7.10	7.10	1	False	-1.21
43277	0.53	0.51	12.75	12.90	4	False	0.12
43277	0.55	0.51	12.55	12.90	4	False	1.09
59321	0.71	0.70	10.79	10.90	1	False	-0.25
59347	0.73	0.72	10.55	10.67	1	False	-0.26
60271	0.21	0.26	16.23	15.57	1	False	0.73
60276	0.84	0.85	9.02	8.85	1	False	-0.88
60299	0.72	0.73	10.67	10.55	1	False	-1.29
60358	0.62	0.63	11.78	11.67	1	False	-1.23
60375	0.50	0.47	13.00	13.30	1	False	0.89
76764	0.84	0.86	9.02	8.68	1	False	-0.02
76778	0.83	0.83	9.18	9.18	1	False	-0.99
76780	0.84	0.85	9.02	8.85	1	False	-0.88
76839	0.66	0.67	11.35	11.24	1	False	-1.30
76845	0.40	0.37	14.01	14.33	1	False	1.06
76863	0.45	0.53	13.50	12.70	1	False	1.76
76871	0.64	0.66	11.57	11.35	1	False	-0.91
76873	0.84	0.85	9.02	8.85	1	False	-0.88
76906	0.67	0.69	11.24	11.02	1	False	-0.84
76916	0.43	0.46	13.71	13.40	1	False	-0.71
76920	0.49	0.48	13.10	13.20	1	False	-0.09
76968	0.47	0.44	13.30	13.60	1	False	0.93
77013	0.96	0.95	6.00	6.42	1	False	0.75
77041	0.57	0.63	12.29	11.67	1	False	1.00

Table K-2. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 4

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
139538	0.69	0.69	11.02	11.02	1	False	-0.88
139542	0.62	0.61	11.78	11.88	1	False	-0.49
139545	0.57	0.58	12.29	12.19	1	False	-1.10
139562	0.77	0.74	10.04	10.43	1	False	0.25
139571	0.46	0.47	13.40	13.30	1	False	-1.02
139630	0.74	0.76	10.43	10.17	1	False	-0.55
139765	0.89	0.91	8.09	7.64	1	False	0.26
139785	0.54	0.56	12.60	12.40	1	False	-0.86
139903	0.26	0.26	15.57	15.57	1	False	-0.55
139911	0.76	0.82	10.17	9.34	1	False	1.29
139915	0.82	0.79	9.34	9.77	1	False	0.36
139934	0.60	0.57	11.99	12.29	1	False	0.15
139950	0.33	0.39	14.76	14.12	1	False	0.35
139956	0.59	0.61	12.09	11.88	1	False	-0.81
139961	0.81	0.80	9.49	9.63	1	False	-0.53
140183	0.54	0.49	12.60	13.10	4	False	0.80
140183	0.53	0.49	12.72	13.10	4	False	0.42
242919	0.80	0.79	9.63	9.77	1	False	-0.54
243107	0.51	0.57	12.90	12.29	1	False	0.37
244352	0.79	0.77	9.77	10.04	1	False	-0.12
248007	0.44	0.59	13.60	12.09	1	True	3.15
248132	0.54	0.63	12.60	11.67	1	False	1.39
43266	0.66	0.63	11.35	11.67	1	False	0.15
62222	0.62	0.63	11.78	11.67	1	False	-1.11
62302	0.75	0.73	10.30	10.55	1	False	-0.16
62317	0.54	0.52	12.60	12.80	1	False	-0.13
62324	0.77	0.76	10.04	10.17	1	False	-0.54
62343	0.69	0.69	11.02	11.02	1	False	-0.88
62355	0.84	0.83	9.02	9.18	1	False	-0.52
76773	0.40	0.45	14.01	13.50	1	False	0.00
76789	0.86	0.83	8.68	9.18	1	False	0.53
76803	0.55	0.53	12.50	12.70	1	False	-0.14
76822	0.74	0.74	10.43	10.43	1	False	-0.92
76832	0.78	0.79	9.91	9.77	1	False	-0.87
76861	0.72	0.67	10.67	11.24	1	False	0.88
76891	0.59	0.74	12.09	10.43	1	True	3.73
76933	0.56	0.55	12.40	12.50	1	False	-0.46
77022	0.69	0.67	11.02	11.24	1	False	-0.18
77024	0.20	0.18	16.37	16.66	1	False	0.43
77026	0.73	0.74	10.55	10.43	1	False	-0.96
77031	0.59	0.62	12.09	11.78	1	False	-0.48
77042	0.44	0.48	13.60	13.20	1	False	-0.31
77046	0.58	0.57	12.19	12.29	1	False	-0.47
77056	0.79	0.84	9.77	9.02	1	False	1.06

Table K-3. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 5

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
140693	0.58	0.57	12.19	12.29	1	False	-0.15
140786	0.86	0.84	8.68	9.02	1	False	0.83
140798	0.57	0.55	12.29	12.50	1	False	0.24
140800	0.33	0.32	14.76	14.87	1	False	-0.15
140823	0.78	0.82	9.91	9.34	1	False	0.32
140829	0.72	0.73	10.67	10.55	1	False	-1.00
140878	0.42	0.47	13.81	13.30	1	False	0.11
140907	0.61	0.63	11.88	11.67	1	False	-1.07
140919	0.47	0.46	13.30	13.40	1	False	-0.17
235941	0.53	0.57	12.70	12.29	1	False	-0.30
236134	0.37	0.36	14.33	14.43	1	False	-0.16
236242	0.62	0.61	11.78	11.88	1	False	-0.14
242883	0.26	0.37	15.57	14.33	1	True	3.02
242922	0.31	0.34	14.98	14.65	1	False	-0.55
242999	0.64	0.75	11.57	10.30	1	True	3.04
243011	0.62	0.60	11.78	11.99	1	False	0.27
43427	0.59	0.56	12.09	12.40	1	False	0.65
43465	0.44	0.45	13.60	13.50	1	False	-0.96
43471	0.70	0.64	10.90	11.57	1	False	2.06
43546	0.53	0.62	12.70	11.78	1	False	1.72
43558	0.77	0.78	10.04	9.91	1	False	-1.04
43566	0.68	0.68	11.13	11.13	1	False	-0.54
43581	0.62	0.70	11.78	10.90	1	False	1.53
43585	0.65	0.65	11.46	11.46	1	False	-0.54
59818	0.53	0.60	12.70	11.99	1	False	0.90
59856	0.58	0.61	12.19	11.88	1	False	-0.68
59861	0.57	0.61	12.29	11.88	1	False	-0.28
59863	0.85	0.86	8.85	8.68	1	False	-1.19
59995	0.73	0.73	10.55	10.55	1	False	-0.53
60383	0.80	0.82	9.63	9.34	1	False	-0.76
60398	0.67	0.69	11.24	11.02	1	False	-1.02
60845	0.74	0.75	10.43	10.30	1	False	-1.01
62035	0.68	0.70	11.13	10.90	1	False	-1.01
77165	0.44	0.49	13.60	13.10	1	False	0.10
77169	0.81	0.80	9.49	9.63	1	False	0.05
77171	0.75	0.77	10.30	10.04	1	False	-0.90
77198	0.46	0.45	13.40	13.50	1	False	-0.17
77205	0.78	0.78	9.91	9.91	1	False	-0.52
77207	0.59	0.56	12.09	12.40	1	False	0.65
77228	0.52	0.52	12.80	12.80	1	False	-0.56
77241	0.63	0.69	11.67	11.02	1	False	0.67
77247	0.71	0.69	10.79	11.02	1	False	0.37
77254	0.31	0.26	14.98	15.57	1	False	1.72
77281	0.79	0.82	9.77	9.34	1	False	-0.21
77286	0.77	0.79	10.04	9.77	1	False	-0.85
77299	0.68	0.73	11.13	10.55	1	False	0.37
77302	0.71	0.74	10.79	10.43	1	False	-0.50
77306	0.61	0.60	11.88	11.99	1	False	-0.14
77309	0.66	0.67	11.35	11.24	1	False	-0.97
77388	0.74	0.74	10.43	10.43	1	False	-0.53

Table K-4. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 6

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
140983	0.59	0.63	12.09	11.67	1	False	0.39
141156	0.31	0.29	14.98	15.21	1	False	-0.06
141165	0.49	0.46	13.10	13.40	1	False	0.11
141175	0.57	0.57	12.29	12.29	1	False	-1.05
141178	0.57	0.59	12.29	12.09	1	False	-0.41
141283	0.79	0.76	9.77	10.17	1	False	0.31
141405	0.49	0.51	13.10	12.90	1	False	-0.47
141406	0.75	0.73	10.30	10.55	1	False	-0.23
141407	0.74	0.73	10.43	10.55	1	False	-0.69
141452	0.65	0.60	11.46	11.99	1	False	0.87
141470	0.63	0.66	11.67	11.35	1	False	0.06
141512	0.69	0.71	11.02	10.79	1	False	-0.25
243207	0.48	0.48	13.20	13.20	1	False	-1.00
43852	0.86	0.88	8.68	8.30	1	False	0.42
43857	0.50	0.49	13.00	13.10	1	False	-0.64
43913	0.52	0.59	12.80	12.09	1	False	1.44
43913	0.57	0.59	12.29	12.09	1	False	-0.41
43944	0.67	0.67	11.24	11.24	1	False	-1.10
43985	0.69	0.67	11.02	11.24	1	False	-0.28
43992	0.52	0.57	12.80	12.29	1	False	0.68
43993	0.47	0.52	13.30	12.80	1	False	0.64
44027	0.39	0.33	14.12	14.76	1	False	1.43
44040	0.64	0.66	11.57	11.35	1	False	-0.33
60883	0.53	0.55	12.70	12.50	1	False	-0.44
60890	0.65	0.64	11.46	11.57	1	False	-0.69
61130	0.86	0.84	8.68	9.02	1	False	0.04
61136	0.46	0.47	13.40	13.30	1	False	-0.85
61151	0.51	0.56	12.90	12.40	1	False	0.67
61162	0.60	0.60	11.99	11.99	1	False	-1.06
62044	0.61	0.54	11.88	12.60	1	False	1.59
62046	0.73	0.70	10.55	10.90	1	False	0.18
62047	0.46	0.45	13.40	13.50	1	False	-0.62
62054	0.72	0.64	10.67	11.57	1	False	2.20
62060	0.85	0.85	8.85	8.85	1	False	-0.99
62062	0.70	0.68	10.90	11.13	1	False	-0.28
62073	0.57	0.56	12.29	12.40	1	False	-0.67
62994	0.45	0.45	13.50	13.50	1	False	-0.99
63054	0.34	0.38	14.65	14.27	4	False	0.10
63054	0.36	0.38	14.46	14.27	4	False	-0.59
77317	0.47	0.48	13.30	13.20	1	False	-0.85
77335	0.48	0.49	13.20	13.10	1	False	-0.84
77337	0.39	0.37	14.12	14.33	1	False	-0.17
77339	0.51	0.49	12.90	13.10	1	False	-0.27
77340	0.62	0.62	11.78	11.78	1	False	-1.07
77376	0.63	0.61	11.67	11.88	1	False	-0.30
77377	0.93	0.93	7.10	7.10	1	False	-0.91
77414	0.73	0.69	10.55	11.02	1	False	0.60
77531	0.62	0.75	11.78	10.30	1	True	4.34
77555	0.77	0.73	10.04	10.55	1	False	0.71
77577	0.61	0.68	11.88	11.13	1	False	1.65
77621	0.76	0.74	10.17	10.43	1	False	-0.22
77646	0.72	0.75	10.67	10.30	1	False	0.27

Table K-5. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 7

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
142365	0.67	0.71	11.24	10.79	1	False	0.15
142368	0.66	0.65	11.35	11.46	1	False	-0.17
142376	0.58	0.56	12.19	12.40	1	False	0.29
142377	0.39	0.39	14.12	14.12	1	False	-0.60
142645	0.73	0.70	10.55	10.90	1	False	0.94
142649	0.66	0.65	11.35	11.46	1	False	-0.17
142680	0.33	0.39	14.76	14.12	1	False	0.93
142735	0.29	0.28	15.21	15.33	1	False	-0.03
142818	0.42	0.37	13.81	14.33	1	False	1.78
43654	0.65	0.65	11.46	11.46	1	False	-0.66
43680	0.30	0.29	15.10	15.21	1	False	-0.04
43719	0.56	0.60	12.40	11.99	1	False	-0.08
43753	0.59	0.57	12.09	12.29	1	False	0.29
43767	0.88	0.84	8.30	9.02	1	False	2.58
43836	0.60	0.58	11.99	12.19	1	False	0.30
43842	0.87	0.91	8.49	7.64	1	False	2.07
43862	0.35	0.43	14.54	13.71	1	False	1.82
43911	0.51	0.52	12.90	12.80	1	False	-1.09
43911	0.50	0.52	13.00	12.80	1	False	-1.06
44173	0.25	0.23	15.70	15.96	1	False	0.63
61244	0.81	0.84	9.49	9.02	1	False	0.25
61279	0.87	0.87	8.49	8.49	1	False	-0.74
61348	0.50	0.54	13.00	12.60	1	False	-0.14
61363	0.54	0.56	12.60	12.40	1	False	-1.04
61374	0.68	0.69	11.13	11.02	1	False	-1.19
61745	0.66	0.65	11.35	11.46	1	False	-0.17
61772	0.90	0.90	7.87	7.87	1	False	-0.75
61781	0.74	0.73	10.43	10.55	1	False	-0.13
61799	0.63	0.66	11.67	11.35	1	False	-0.47
61874	0.61	0.65	11.88	11.46	1	False	-0.01
77389	0.39	0.41	14.12	13.91	1	False	-1.06
86264	0.45	0.48	13.50	13.20	1	False	-0.61
86288	0.60	0.61	11.99	11.88	1	False	-1.13
86290	0.79	0.80	9.77	9.63	1	False	-1.25
86302	0.51	0.52	12.90	12.80	1	False	-1.09
86330	0.60	0.57	11.99	12.29	1	False	0.77
86333	0.61	0.59	11.88	12.09	1	False	0.30
86339	0.49	0.54	13.10	12.60	1	False	0.32
86382	0.79	0.79	9.77	9.77	1	False	-0.70
86431	0.42	0.42	13.81	13.81	1	False	-0.60
86445	0.39	0.38	14.12	14.22	1	False	-0.11
86535	0.79	0.79	9.77	9.77	1	False	-0.70
86549	0.52	0.55	12.80	12.50	1	False	-0.59
86568	0.49	0.52	13.10	12.80	1	False	-0.60
86574	0.59	0.62	12.09	11.78	1	False	-0.53
86597	0.46	0.56	13.40	12.40	1	False	2.63
86606	0.53	0.56	12.70	12.40	1	False	-0.58
86672	0.69	0.72	11.02	10.67	1	False	-0.33
86681	0.80	0.85	9.63	8.85	1	False	1.68
86687	0.46	0.47	13.40	13.30	1	False	-1.08
86700	0.25	0.32	15.70	14.87	1	False	1.75

Table K-6. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 8

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
144191	0.67	0.68	11.24	11.13	1	False	-0.68
144436	0.48	0.47	13.20	13.30	1	False	-0.18
144499	0.68	0.62	11.13	11.78	1	False	1.29
144905	0.78	0.76	9.91	10.17	1	False	-0.33
144942	0.59	0.54	12.09	12.60	1	False	1.00
144955	0.69	0.70	11.02	10.90	1	False	-0.62
166323	0.47	0.55	13.30	12.50	1	False	1.30
243699	0.51	0.51	12.90	12.90	1	False	-0.60
243699	0.49	0.51	13.10	12.90	1	False	-0.77
243710	0.40	0.43	14.01	13.71	1	False	-0.60
243710	0.40	0.43	14.01	13.71	1	False	-0.60
243715	0.46	0.58	13.40	12.19	1	False	2.70
243715	0.54	0.58	12.60	12.19	1	False	0.06
244461	0.59	0.59	12.09	12.09	1	False	-0.78
244462	0.46	0.42	13.40	13.81	1	False	0.93
244684	0.44	0.50	13.60	13.00	1	False	0.53
35019	0.88	0.90	8.30	7.87	1	False	1.07
35025	0.69	0.72	11.02	10.67	1	False	0.20
43840	0.34	0.38	14.65	14.22	1	False	-0.31
43888	0.53	0.59	12.70	12.09	1	False	0.75
44154	0.50	0.49	13.00	13.10	1	False	-0.23
44181	0.35	0.53	14.54	12.70	1	True	4.67
44209	0.65	0.62	11.46	11.78	1	False	0.20
44210	0.65	0.63	11.46	11.67	1	False	-0.17
44213	0.29	0.30	15.21	15.10	1	False	-0.50
61198	0.40	0.38	14.01	14.22	1	False	0.37
63038	0.82	0.81	9.34	9.49	1	False	-0.85
63048	0.73	0.73	10.55	10.55	1	False	-0.92
63066	0.38	0.39	14.22	14.12	1	False	-0.68
63106	0.85	0.82	8.85	9.34	1	False	0.21
63115	0.46	0.46	13.40	13.40	1	False	-0.49
63239	0.56	0.56	12.40	12.40	1	False	-0.71
63242	0.62	0.62	11.78	11.78	1	False	-0.85
63256	0.72	0.69	10.67	11.02	1	False	0.13
86542	0.72	0.67	10.67	11.24	1	False	0.91
87566	0.34	0.35	14.65	14.54	1	False	-0.60
87571	0.58	0.58	12.19	12.19	1	False	-0.76
87576	0.46	0.45	13.40	13.50	1	False	-0.14
87653	0.54	0.49	12.60	13.10	1	False	1.09
87874	0.51	0.50	12.90	12.97	4	False	-0.34
87874	0.52	0.50	12.80	12.97	4	False	-0.01
88026	0.75	0.79	10.30	9.77	1	False	0.99
88086	0.52	0.52	12.80	12.80	1	False	-0.62
88177	0.76	0.76	10.17	10.17	1	False	-0.84
88187	0.36	0.41	14.43	13.91	1	False	0.07
88216	0.72	0.72	10.67	10.67	1	False	-0.94
88277	0.42	0.43	13.81	13.71	1	False	-0.76
88313	0.53	0.51	12.70	12.90	1	False	0.06
88331	0.47	0.51	13.30	12.90	1	False	-0.11
88352	0.69	0.70	11.02	10.90	1	False	-0.62
88361	0.53	0.53	12.70	12.70	1	False	-0.65
88838	0.64	0.63	11.57	11.67	1	False	-0.52
88892	0.51	0.53	12.90	12.70	1	False	-0.73

Table K-7. 2010–11 MontCAS: Delta Analysis Results – Mathematics Grade 10

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
144377	0.61	0.63	11.88	11.67	1	False	-0.45
144570	0.45	0.46	13.50	13.40	1	False	-0.98
144825	0.56	0.56	12.40	12.40	1	False	-0.84
144841	0.54	0.54	12.60	12.60	1	False	-0.83
144904	0.72	0.70	10.67	10.90	1	False	-0.16
144996	0.49	0.50	13.10	13.00	1	False	-0.95
145015	0.67	0.68	11.24	11.13	1	False	-0.75
145022	0.53	0.50	12.70	13.00	1	False	0.25
145032	0.48	0.52	13.20	12.80	1	False	0.11
145292	0.56	0.57	12.40	12.29	1	False	-0.88
145637	0.60	0.58	11.99	12.19	1	False	-0.15
145719	0.50	0.53	13.00	12.70	1	False	-0.23
146070	0.56	0.57	12.40	12.29	1	False	-0.88
146535	0.52	0.49	12.80	13.10	1	False	0.26
146539	0.54	0.51	12.60	12.90	1	False	0.24
146556	0.45	0.42	13.50	13.81	1	False	0.33
146572	0.53	0.51	12.70	12.90	1	False	-0.10
146573	0.55	0.39	12.50	14.12	1	True	4.91
146580	0.52	0.56	12.80	12.40	1	False	0.15
243050	0.55	0.61	12.50	11.88	1	False	0.93
248806	0.35	0.38	14.54	14.22	1	False	-0.29
34804	0.51	0.52	12.90	12.80	1	False	-0.93
43631	0.26	0.26	15.57	15.57	1	False	-0.57
43717	0.70	0.73	10.90	10.55	1	False	0.14
43725	0.26	0.26	15.57	15.57	1	False	-0.57
43951	0.30	0.32	15.10	14.87	1	False	-0.67
43969	0.88	0.87	8.30	8.49	1	False	-0.50
59403	0.65	0.65	11.46	11.46	1	False	-0.92
59407	0.30	0.35	15.13	14.57	4	False	0.50
61301	0.49	0.50	13.10	13.00	1	False	-0.95
61305	0.26	0.30	15.57	15.10	1	False	0.17
61324	0.88	0.87	8.30	8.49	1	False	-0.50
62279	0.66	0.70	11.35	10.90	1	False	0.43
62319	0.44	0.48	13.60	13.20	1	False	0.08
62361	0.57	0.62	12.29	11.78	1	False	0.60
62365	0.43	0.42	13.71	13.81	1	False	-0.37
62391	0.22	0.25	16.16	15.73	4	False	-0.05
77358	0.59	0.57	12.09	12.29	1	False	-0.14
77364	0.84	0.85	9.02	8.85	1	False	-0.36
77369	0.60	0.64	11.99	11.57	1	False	0.28
77374	0.29	0.27	15.21	15.45	1	False	0.24
77420	0.32	0.39	14.87	14.12	1	False	1.22
77428	0.65	0.62	11.46	11.78	1	False	0.21
77430	0.50	0.49	13.00	13.10	1	False	-0.43
77432	0.68	0.63	11.13	11.67	1	False	0.98
77501	0.37	0.38	14.33	14.22	1	False	-1.03
77503	0.43	0.41	13.71	13.91	1	False	0.00
77512	0.25	0.34	15.70	14.65	1	False	2.19
77544	0.54	0.61	12.60	11.88	1	False	1.28
77546	0.43	0.44	13.71	13.60	1	False	-1.00
77592	0.42	0.44	13.81	13.60	1	False	-0.64
77602	0.21	0.16	16.23	16.98	1	False	2.15
77629	0.42	0.41	13.81	13.91	1	False	-0.36
77644	0.22	0.25	16.09	15.70	1	False	-0.17

Table K-8. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 3

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
150634	0.42	0.46	13.81	13.40	1	False	-0.20
150644	0.57	0.59	12.29	12.09	1	False	-0.99
150648	0.83	0.83	9.18	9.18	1	False	-1.21
150653	0.90	0.88	7.87	8.30	1	False	0.64
150656	0.58	0.60	12.19	11.99	1	False	-0.98
150664	0.84	0.86	9.02	8.68	1	False	-0.07
150670	0.77	0.78	10.04	9.91	1	False	-1.13
150711	0.60	0.66	11.99	11.35	1	False	1.02
150713	0.58	0.56	12.19	12.40	1	False	-0.01
150718	0.84	0.85	9.02	8.85	1	False	-0.88
150721	0.88	0.88	8.30	8.30	1	False	-1.29
150725	0.92	0.89	7.38	8.09	1	False	1.92
150725	0.92	0.89	7.38	8.09	1	False	1.92
150869	0.91	0.93	7.64	7.10	1	False	0.96
150951	0.86	0.89	8.68	8.09	1	False	1.07
150953	0.72	0.77	10.67	10.04	1	False	1.08
150962	0.82	0.81	9.34	9.49	1	False	-0.51
150969	0.76	0.75	10.17	10.30	1	False	-0.54
150971	0.91	0.90	7.64	7.87	1	False	-0.26
150974	0.73	0.72	10.55	10.67	1	False	-0.54
150975	0.92	0.91	7.38	7.64	1	False	-0.18
150977	0.72	0.71	10.67	10.79	1	False	-0.54
150978	0.90	0.90	7.87	7.87	1	False	-1.32
150982	0.43	0.50	13.71	13.00	1	False	1.19
150983	0.80	0.80	9.63	9.63	1	False	-1.17
150986	0.67	0.66	11.24	11.35	1	False	-0.53
150988	0.52	0.48	12.80	13.20	1	False	0.95
150991	0.44	0.41	13.65	13.94	4	False	0.47
151127	0.89	0.92	8.09	7.38	1	False	1.72
151131	0.61	0.60	11.88	11.99	1	False	-0.50
151133	0.73	0.70	10.55	10.90	1	False	0.54
151135	0.71	0.74	10.79	10.43	1	False	-0.15
151141	0.55	0.58	12.50	12.19	1	False	-0.55
151156	0.79	0.83	9.77	9.18	1	False	1.00
151166	0.51	0.48	12.90	13.20	1	False	0.50
151173	0.71	0.69	10.79	11.02	1	False	-0.01
151176	0.86	0.87	8.68	8.49	1	False	-0.77
151186	0.74	0.76	10.43	10.17	1	False	-0.61
151193	0.56	0.55	12.40	12.50	1	False	-0.46
151194	0.48	0.50	13.20	13.00	1	False	-1.09
151200	0.79	0.80	9.77	9.63	1	False	-1.07
151203	0.69	0.77	11.02	10.04	1	False	2.65
151207	0.42	0.48	13.81	13.20	1	False	0.72
151212	0.62	0.65	11.78	11.46	1	False	-0.42
151215	0.80	0.81	9.63	9.49	1	False	-1.04
151227	0.59	0.57	12.09	12.29	1	False	-0.01
151240	0.35	0.35	14.57	14.51	4	False	-0.99
153156	0.61	0.57	11.88	12.29	1	False	0.92
67148	0.70	0.74	10.90	10.43	1	False	0.37
67151	0.71	0.78	10.79	9.91	1	False	2.22
67155	0.88	0.87	8.30	8.49	1	False	-0.39
67167	0.81	0.78	9.49	9.91	1	False	0.76
67184	0.69	0.65	11.02	11.46	1	False	0.98
67193	0.75	0.73	10.30	10.55	1	False	0.02
67193	0.72	0.73	10.67	10.55	1	False	-1.24

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
67198	0.48	0.49	13.20	13.10	1	False	-1.32
67198	0.46	0.49	13.40	13.10	1	False	-0.65

Table K-9. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 4

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
151597	0.70	0.69	10.90	11.02	1	False	-0.11
151597	0.68	0.69	11.13	11.02	1	False	-0.83
151612	0.83	0.83	9.18	9.18	1	False	0.05
151615	0.83	0.84	9.18	9.02	1	False	-0.42
151621	0.64	0.60	11.57	11.99	1	False	0.59
151626	0.83	0.84	9.18	9.02	1	False	-0.42
151632	0.52	0.52	12.80	12.80	1	False	-0.98
151635	0.68	0.71	11.13	10.79	1	False	-0.51
151637	0.86	0.83	8.68	9.18	1	False	1.65
151638	0.69	0.73	11.02	10.55	1	False	-0.18
151639	0.64	0.62	11.57	11.78	1	False	-0.01
151644	0.72	0.68	10.67	11.13	1	False	0.96
151655	0.60	0.60	11.99	11.99	1	False	-0.75
151659	0.36	0.34	14.43	14.62	4	False	-0.89
151672	0.84	0.87	9.02	8.49	1	False	-0.57
151679	0.58	0.56	12.19	12.40	1	False	-0.21
151683	0.81	0.78	9.49	9.91	1	False	1.19
151684	0.58	0.63	12.19	11.67	1	False	0.30
151686	0.65	0.69	11.46	11.02	1	False	-0.13
151695	0.74	0.79	10.43	9.77	1	False	0.18
151697	0.79	0.83	9.77	9.18	1	False	-0.18
151740	0.72	0.80	10.67	9.63	1	False	1.36
151743	0.86	0.88	8.68	8.30	1	False	-0.91
151745	0.69	0.59	11.02	12.09	1	False	2.64
151751	0.74	0.75	10.43	10.30	1	False	-0.67
151762	0.63	0.58	11.67	12.19	1	False	0.85
151772	0.41	0.45	13.94	13.50	4	False	0.55
151780	0.41	0.40	13.91	14.01	1	False	-0.99
151782	0.88	0.88	8.30	8.30	1	False	0.30
151783	0.74	0.86	10.43	8.68	1	True	3.36
151788	0.69	0.68	11.02	11.13	1	False	-0.15
151790	0.71	0.69	10.79	11.02	1	False	0.26
151794	0.63	0.63	11.67	11.67	1	False	-0.66
151800	0.80	0.76	9.63	10.17	1	False	1.49
66546	0.53	0.51	12.70	12.90	1	False	-0.37
66546	0.53	0.51	12.70	12.90	1	False	-0.37
66550	0.82	0.83	9.34	9.18	1	False	-0.45
66555	0.56	0.59	12.40	12.09	1	False	-0.26
66568	0.75	0.78	10.30	9.91	1	False	-0.61
66571	0.71	0.71	10.79	10.79	1	False	-0.41
66578	0.48	0.48	13.20	13.20	1	False	-0.92
66586	0.69	0.70	11.02	10.90	1	False	-0.80
66597	0.81	0.84	9.49	9.02	1	False	-0.62
66597	0.81	0.84	9.49	9.02	1	False	-0.62
66612	0.65	0.69	11.46	11.02	1	False	-0.13
66612	0.69	0.69	11.02	11.02	1	False	-0.47
66615	0.85	0.87	8.85	8.49	1	False	-0.90
66615	0.85	0.87	8.85	8.49	1	False	-0.90

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
66627	0.79	0.83	9.77	9.18	1	False	-0.18
66629	0.70	0.77	10.90	10.04	1	False	0.92
66634	0.56	0.70	12.40	10.90	1	True	3.18
66650	0.82	0.84	9.34	9.02	1	False	-0.91
67140	0.67	0.73	11.24	10.55	1	False	0.53
67149	0.81	0.88	9.49	8.30	1	False	1.47
67154	0.82	0.85	9.34	8.85	1	False	-0.61
67158	0.57	0.59	12.29	12.09	1	False	-0.58
67160	0.77	0.79	10.04	9.77	1	False	-0.98
67164	0.73	0.77	10.55	10.04	1	False	-0.21
67166	0.78	0.80	9.91	9.63	1	False	-0.96

Table K-10. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 5

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
150413	0.62	0.66	11.78	11.35	1	False	-0.37
150415	0.83	0.85	9.18	8.85	1	False	-0.66
150417	0.81	0.83	9.49	9.18	1	False	-0.76
150419	0.75	0.74	10.30	10.43	1	False	0.06
150423	0.72	0.73	10.67	10.55	1	False	-0.85
150425	0.75	0.81	10.30	9.49	1	False	1.12
150427	0.87	0.89	8.49	8.09	1	False	-0.38
150432	0.51	0.51	12.90	12.90	1	False	-0.33
150440	0.89	0.90	8.09	7.87	1	False	-1.04
150442	0.86	0.77	8.68	10.04	1	True	4.66
150450	0.39	0.42	14.12	13.81	1	False	-0.88
150456	0.45	0.47	13.55	13.35	4	False	-1.07
150470	0.90	0.90	7.87	7.87	1	False	-0.48
150470	0.87	0.90	8.49	7.87	1	False	0.45
150474	0.74	0.77	10.43	10.04	1	False	-0.50
150480	0.86	0.88	8.68	8.30	1	False	-0.46
150483	0.77	0.82	10.04	9.34	1	False	0.72
150485	0.76	0.74	10.17	10.43	1	False	0.53
150488	0.88	0.86	8.30	8.68	1	False	0.95
150489	0.66	0.72	11.35	10.67	1	False	0.59
150491	0.81	0.79	9.49	9.77	1	False	0.64
150492	0.79	0.81	9.77	9.49	1	False	-0.84
150492	0.83	0.81	9.18	9.49	1	False	0.70
150493	0.80	0.80	9.63	9.63	1	False	-0.43
150494	0.78	0.80	9.91	9.63	1	False	-0.88
150505	0.62	0.67	11.78	11.24	1	False	0.04
150518	0.49	0.50	13.08	13.03	4	False	-0.52
150527	0.61	0.62	11.88	11.78	1	False	-0.76
150530	0.84	0.88	9.02	8.30	1	False	0.81
150536	0.71	0.65	10.79	11.46	1	False	2.12
150547	0.82	0.80	9.34	9.63	1	False	0.67
150548	0.52	0.56	12.80	12.40	1	False	-0.49
150551	0.73	0.68	10.55	11.13	1	False	1.77
150578	0.64	0.69	11.57	11.02	1	False	0.09
150578	0.67	0.69	11.24	11.02	1	False	-1.12
150580	0.86	0.90	8.68	7.87	1	False	1.13
150586	0.75	0.78	10.30	9.91	1	False	-0.47
150586	0.73	0.78	10.55	9.91	1	False	0.45
150591	0.83	0.86	9.18	8.68	1	False	-0.01

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
150593	0.65	0.68	11.46	11.13	1	False	-0.73
150594	0.76	0.79	10.17	9.77	1	False	-0.43
150596	0.62	0.58	11.78	12.19	1	False	1.19
155431	0.75	0.74	10.30	10.43	1	False	0.06
41321	0.70	0.70	10.90	10.90	1	False	-0.39
41322	0.41	0.46	13.91	13.40	1	False	-0.13
41323	0.71	0.74	10.79	10.43	1	False	-0.60
41326	0.78	0.81	9.91	9.49	1	False	-0.33
41328	0.64	0.69	11.57	11.02	1	False	0.09
41329	0.59	0.63	12.09	11.67	1	False	-0.42
41331	0.80	0.82	9.63	9.34	1	False	-0.80
65580	0.54	0.54	12.60	12.60	1	False	-0.34
65593	0.80	0.78	9.63	9.91	1	False	0.61
65599	0.63	0.65	11.67	11.46	1	False	-1.17
65607	0.67	0.69	11.24	11.02	1	False	-1.12
65611	0.54	0.63	12.60	11.67	1	False	1.47
65613	0.56	0.60	12.40	11.99	1	False	-0.46
65616	0.82	0.84	9.34	9.02	1	False	-0.72

Table K-11. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 6

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
151273	0.85	0.85	8.85	8.85	1	False	0.55
151274	0.67	0.65	11.24	11.46	1	False	0.38
151278	0.73	0.75	10.55	10.30	1	False	-0.77
151282	0.82	0.85	9.34	8.85	1	False	-1.06
151295	0.69	0.71	11.02	10.79	1	False	-0.89
151296	0.83	0.88	9.18	8.30	1	False	-0.40
151302	0.46	0.52	13.40	12.80	1	False	0.23
151312	0.62	0.61	11.78	11.88	1	False	-0.15
151315	0.72	0.72	10.67	10.67	1	False	-0.08
151321	0.64	0.60	11.57	11.99	1	False	0.87
151321	0.64	0.60	11.57	11.99	1	False	0.87
151325	0.64	0.60	11.57	11.99	1	False	0.87
151337	0.82	0.79	9.34	9.77	1	False	1.69
151338	0.50	0.52	13.00	12.80	1	False	-1.11
151341	0.66	0.73	11.35	10.55	1	False	0.11
42099	0.68	0.81	11.13	9.49	1	False	2.55
42100	0.86	0.91	8.68	7.64	1	False	-0.10
42102	0.74	0.76	10.43	10.17	1	False	-0.75
42104	0.83	0.86	9.18	8.68	1	False	-1.07
42106	0.46	0.61	13.40	11.88	1	False	2.98
42118	0.84	0.90	9.02	7.87	1	False	0.34
42130	0.85	0.92	8.85	7.38	1	False	1.26
42131	0.83	0.82	9.18	9.34	1	False	0.90
42132	0.81	0.77	9.49	10.04	1	False	2.00
42282	0.68	0.66	11.13	11.35	1	False	0.42
42284	0.66	0.71	11.35	10.79	1	False	-0.60
42287	0.83	0.90	9.18	7.87	1	False	0.88
42287	0.86	0.90	8.68	7.87	1	False	-0.81
68240	0.50	0.53	13.00	12.70	1	False	-0.81
68242	0.60	0.62	11.99	11.78	1	False	-1.16
68247	0.66	0.69	11.35	11.02	1	False	-1.29
68250	0.47	0.45	13.30	13.50	1	False	-0.39

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
68260	0.70	0.70	10.90	10.90	1	False	-0.16
68260	0.65	0.70	11.46	10.90	1	False	-0.58
68263	0.61	0.66	11.88	11.35	1	False	-0.51
68265	0.69	0.71	11.02	10.79	1	False	-0.89
68265	0.70	0.71	10.90	10.79	1	False	-0.51
68269	0.46	0.47	13.40	13.30	1	False	-1.27
68283	0.73	0.75	10.55	10.30	1	False	-0.77
68285	0.64	0.68	11.57	11.13	1	False	-0.90
68291	0.56	0.46	12.40	13.40	1	False	2.33
68294	0.36	0.41	14.49	13.94	4	False	0.46
95414	0.80	0.85	9.63	8.85	1	False	-0.55
95417	0.66	0.75	11.35	10.30	1	False	0.85
95422	0.81	0.86	9.49	8.68	1	False	-0.51
95426	0.44	0.48	13.60	13.20	1	False	-0.29
95429	0.87	0.88	8.49	8.30	1	False	0.10
95446	0.65	0.71	11.46	10.79	1	False	-0.24
95447	0.86	0.87	8.68	8.49	1	False	0.06
95448	0.77	0.80	10.04	9.63	1	False	-1.09
95451	0.58	0.65	12.19	11.46	1	False	0.20
95451	0.57	0.65	12.29	11.46	1	False	0.54
95455	0.70	0.78	10.90	9.91	1	False	0.52
95455	0.73	0.78	10.55	9.91	1	False	-0.66
95456	0.51	0.46	12.90	13.40	1	False	0.65
95459	0.66	0.69	11.35	11.02	1	False	-1.29
95469	0.44	0.45	13.65	13.48	4	False	-0.95

Table K-12. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 7

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
148792	0.75	0.76	10.30	10.17	1	False	-1.21
148807	0.62	0.69	11.78	11.02	1	False	1.15
148811	0.61	0.66	11.88	11.35	1	False	0.19
148827	0.62	0.66	11.78	11.35	1	False	-0.24
148834	0.81	0.82	9.49	9.34	1	False	-1.30
148839	0.85	0.90	8.85	7.87	1	False	2.17
148843	0.67	0.71	11.24	10.79	1	False	-0.11
148849	0.74	0.73	10.43	10.55	1	False	-0.17
148854	0.74	0.83	10.43	9.18	1	True	3.20
148875	0.76	0.81	10.17	9.49	1	False	0.90
148883	0.71	0.79	10.79	9.77	1	False	2.23
148887	0.48	0.53	13.18	12.67	4	False	0.01
148965	0.77	0.76	10.04	10.17	1	False	-0.15
148970	0.76	0.78	10.17	9.91	1	False	-0.86
148970	0.79	0.78	9.77	9.91	1	False	-0.13
148972	0.83	0.82	9.18	9.34	1	False	-0.08
148973	0.85	0.84	8.85	9.02	1	False	-0.04
148977	0.66	0.65	11.35	11.46	1	False	-0.18
148982	0.57	0.58	12.29	12.19	1	False	-1.02
148982	0.59	0.58	12.09	12.19	1	False	-0.18
148988	0.81	0.78	9.49	9.91	1	False	1.04
148993	0.50	0.48	13.00	13.20	1	False	0.26
149003	0.75	0.72	10.30	10.67	1	False	0.84
149007	0.81	0.83	9.49	9.18	1	False	-0.66
149008	0.64	0.68	11.57	11.13	1	False	-0.19

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
149016	0.48	0.49	13.20	13.10	4	False	-0.98
149061	0.72	0.68	10.67	11.13	1	False	1.25
149064	0.81	0.79	9.49	9.77	1	False	0.48
149066	0.77	0.85	10.04	8.85	1	False	3.00
149080	0.75	0.77	10.30	10.04	1	False	-0.89
149083	0.73	0.76	10.55	10.17	1	False	-0.41
149083	0.74	0.76	10.43	10.17	1	False	-0.91
149088	0.74	0.72	10.43	10.67	1	False	0.33
149089	0.79	0.79	9.77	9.77	1	False	-0.70
149089	0.81	0.79	9.49	9.77	1	False	0.48
149090	0.74	0.76	10.43	10.17	1	False	-0.91
149091	0.80	0.80	9.63	9.63	1	False	-0.70
149093	0.55	0.61	12.50	11.88	1	False	0.51
149098	0.60	0.59	11.99	12.09	1	False	-0.18
154487	0.83	0.86	9.18	8.68	1	False	0.19
92359	0.76	0.80	10.17	9.63	1	False	0.30
92384	0.92	0.91	7.38	7.64	1	False	0.27
92391	0.87	0.84	8.49	9.02	1	False	1.44
92395	0.69	0.69	11.02	11.02	1	False	-0.65
92397	0.84	0.85	9.02	8.85	1	False	-1.21
92399	0.73	0.75	10.55	10.30	1	False	-0.94
92404	0.77	0.80	10.04	9.63	1	False	-0.24
92404	0.81	0.80	9.49	9.63	1	False	-0.11
92453	0.53	0.56	12.70	12.40	1	False	-0.80
92453	0.53	0.56	12.70	12.40	1	False	-0.80
92455	0.86	0.86	8.68	8.68	1	False	-0.74
92455	0.86	0.86	8.68	8.68	1	False	-0.74
92458	0.70	0.71	10.90	10.79	1	False	-1.13
92458	0.71	0.71	10.79	10.79	1	False	-0.66
92462	0.68	0.63	11.13	11.67	1	False	1.61
92462	0.67	0.63	11.24	11.67	1	False	1.15
92464	0.75	0.78	10.30	9.91	1	False	-0.33
92464	0.77	0.78	10.04	9.91	1	False	-1.24
92471	0.80	0.80	9.63	9.63	1	False	-0.70
92471	0.81	0.80	9.49	9.63	1	False	-0.11
92472	0.71	0.75	10.79	10.30	1	False	0.04
92472	0.72	0.75	10.67	10.30	1	False	-0.45

Table K-13. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 8

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
149129	0.67	0.67	11.24	11.24	1	False	-0.99
149130	0.53	0.55	12.70	12.50	1	False	-0.72
149134	0.70	0.73	10.90	10.55	1	False	0.09
149136	0.61	0.63	11.88	11.67	1	False	-0.64
149139	0.74	0.76	10.43	10.17	1	False	-0.37
149144	0.76	0.81	10.17	9.49	1	False	1.69
149165	0.66	0.60	11.35	11.99	1	False	2.01
149170	0.84	0.84	9.02	9.02	1	False	-1.11
149170	0.84	0.84	9.02	9.02	1	False	-1.11
149171	0.86	0.86	8.68	8.68	1	False	-1.13
149171	0.85	0.86	8.85	8.68	1	False	-0.64
149174	0.60	0.62	11.99	11.78	1	False	-0.65
149176	0.73	0.73	10.55	10.55	1	False	-1.03

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
149179	0.86	0.84	8.68	9.02	1	False	0.49
149181	0.77	0.78	10.04	9.91	1	False	-0.90
149187	0.78	0.77	9.91	10.04	1	False	-0.44
149188	0.77	0.76	10.04	10.17	1	False	-0.44
149190	0.73	0.73	10.55	10.55	1	False	-1.03
149191	0.79	0.77	9.77	10.04	1	False	0.20
149193	0.57	0.59	12.29	12.14	4	False	-0.93
149261	0.73	0.71	10.55	10.79	1	False	0.09
149262	0.71	0.62	10.79	11.78	1	True	3.65
149269	0.84	0.85	9.02	8.85	1	False	-0.69
149276	0.71	0.69	10.79	11.02	1	False	0.07
149281	0.64	0.61	11.57	11.88	1	False	0.51
149287	0.67	0.69	11.24	11.02	1	False	-0.54
149326	0.69	0.67	11.02	11.24	1	False	0.05
149328	0.85	0.84	8.85	9.02	1	False	-0.33
149331	0.74	0.75	10.43	10.30	1	False	-0.97
149335	0.92	0.90	7.38	7.87	1	False	1.13
149335	0.91	0.90	7.64	7.87	1	False	-0.07
149339	0.85	0.80	8.85	9.63	1	False	2.55
149340	0.73	0.75	10.55	10.30	1	False	-0.40
149341	0.67	0.66	11.24	11.35	1	False	-0.48
149342	0.57	0.59	12.29	12.09	1	False	-0.69
149344	0.72	0.76	10.67	10.17	1	False	0.76
149347	0.63	0.65	11.67	11.46	1	False	-0.61
149349	0.61	0.61	11.88	11.88	1	False	-0.96
149352	0.76	0.79	10.17	9.77	1	False	0.35
149353	0.73	0.74	10.55	10.43	1	False	-0.98
149354	0.66	0.62	11.35	11.78	1	False	1.03
149355	0.64	0.61	11.57	11.88	1	False	0.51
149359	0.75	0.78	10.30	9.91	1	False	0.29
149360	0.75	0.80	10.30	9.63	1	False	1.60
149362	0.88	0.91	8.30	7.64	1	False	1.68
149368	0.54	0.58	12.62	12.24	4	False	0.12
149372	0.51	0.55	12.90	12.50	1	False	0.21
149372	0.52	0.55	12.80	12.50	1	False	-0.25
149377	0.68	0.65	11.13	11.46	1	False	0.55
149379	0.78	0.82	9.91	9.34	1	False	1.17
149380	0.82	0.80	9.34	9.63	1	False	0.29
149383	0.71	0.75	10.79	10.30	1	False	0.71
149388	0.66	0.68	11.35	11.13	1	False	-0.56
152410	0.75	0.76	10.30	10.17	1	False	-0.95
152841	0.74	0.77	10.43	10.04	1	False	0.25
153158	0.65	0.68	11.46	11.13	1	False	-0.06
153228	0.84	0.84	9.02	9.02	1	False	-1.11
153228	0.81	0.84	9.49	9.02	1	False	0.69
156089	0.62	0.62	11.78	11.78	1	False	-0.96

Table K-14. 2010–11 MontCAS: Delta Analysis Results – Reading Grade 10

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
149467	0.80	0.87	9.63	8.49	1	False	1.63
149468	0.62	0.68	11.78	11.13	1	False	0.56
149471	0.65	0.73	11.46	10.55	1	False	1.32
149472	0.68	0.76	11.13	10.17	1	False	1.38
149482	0.83	0.85	9.18	8.85	1	False	-1.04
149483	0.71	0.79	10.79	9.77	1	False	1.49
149486	0.82	0.85	9.34	8.85	1	False	-0.52
149515	0.69	0.69	11.02	11.02	1	False	-0.77
149521	0.74	0.76	10.43	10.17	1	False	-1.01
149525	0.76	0.74	10.17	10.43	1	False	0.22
149528	0.77	0.74	10.04	10.43	1	False	0.67
149537	0.85	0.83	8.85	9.18	1	False	0.77
149538	0.77	0.79	10.04	9.77	1	False	-1.04
149541	0.86	0.85	8.68	8.85	1	False	0.32
149544	0.84	0.85	9.02	8.85	1	False	-0.85
149545	0.71	0.72	10.79	10.67	1	False	-1.09
149548	0.82	0.83	9.34	9.18	1	False	-0.88
149549	0.71	0.70	10.79	10.90	1	False	-0.35
149550	0.71	0.73	10.79	10.55	1	False	-0.98
149551	0.58	0.56	12.19	12.40	1	False	-0.38
149551	0.58	0.56	12.19	12.40	1	False	-0.38
149554	0.75	0.77	10.30	10.04	1	False	-1.02
149555	0.81	0.83	9.49	9.18	1	False	-1.05
149556	0.82	0.81	9.34	9.49	1	False	0.09
149558	0.67	0.71	11.24	10.79	1	False	-0.19
149560	0.66	0.66	11.35	11.35	1	False	-0.84
149564	0.72	0.76	10.67	10.17	1	False	-0.19
149566	0.47	0.51	13.33	12.90	4	False	0.20
149610	0.59	0.57	12.09	12.29	1	False	-0.36
149611	0.83	0.84	9.18	9.02	1	False	-0.87
149613	0.50	0.46	13.00	13.40	1	False	0.07
149615	0.75	0.69	10.30	11.02	1	False	1.67
149616	0.56	0.65	12.40	11.46	1	False	1.61
149623	0.51	0.50	12.90	13.00	1	False	-0.87
149627	0.82	0.82	9.34	9.34	1	False	-0.39
149627	0.83	0.82	9.18	9.34	1	False	0.14
149629	0.75	0.76	10.30	10.17	1	False	-1.01
149630	0.53	0.47	12.70	13.30	1	False	0.77
149631	0.73	0.70	10.55	10.90	1	False	0.46
149633	0.71	0.63	10.79	11.67	1	False	2.10
149638	0.64	0.64	11.57	11.57	1	False	-0.89
149648	0.44	0.47	13.60	13.33	4	False	-0.22
149869	0.60	0.62	11.99	11.78	1	False	-0.80
149883	0.83	0.86	9.18	8.68	1	False	-0.49
149887	0.69	0.59	11.02	12.09	1	False	2.65
149894	0.71	0.73	10.79	10.55	1	False	-0.98
149896	0.72	0.78	10.67	9.91	1	False	0.65
149905	0.77	0.80	10.04	9.63	1	False	-0.59
149950	0.90	0.90	7.87	7.87	1	False	-0.06
42224	0.69	0.67	11.02	11.24	1	False	-0.06
42225	0.72	0.73	10.67	10.55	1	False	-1.07
42226	0.73	0.79	10.55	9.77	1	False	0.68
42228	0.71	0.74	10.79	10.43	1	False	-0.59
42230	0.64	0.65	11.57	11.46	1	False	-1.21
42232	0.63	0.72	11.67	10.67	1	False	1.66
42233	0.71	0.80	10.79	9.63	1	False	1.94

Table K-15. 2010–11 MontCAS: Delta Analysis Results – Science Grade 4

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
119971	0.85	0.86	8.85	8.68	8.77	False	-0.81
120003	0.66	0.66	11.35	11.35	11.47	False	-0.65
120014	0.71	0.72	10.79	10.67	10.78	False	-1.22
120024	0.48	0.50	13.20	13.00	13.13	False	-0.92
120106	0.40	0.39	14.01	14.12	14.26	False	0.04
120162	0.65	0.63	11.46	11.67	11.79	False	0.47
120312	0.50	0.47	13.00	13.30	13.44	False	1.01
120312	0.51	0.47	12.92	13.30	13.44	False	1.40
120617	0.82	0.83	9.34	9.18	9.28	False	-0.94
134675	0.67	0.68	11.24	11.13	11.24	False	-1.24
134745	0.45	0.49	13.50	13.10	13.24	False	0.12
142358	0.51	0.52	12.90	12.80	12.93	False	-1.09
55442	0.55	0.55	12.50	12.50	12.63	False	-0.59
55442	0.53	0.55	12.70	12.50	12.63	False	-0.88
56225	0.42	0.50	13.81	13.00	13.13	False	2.23
56225	0.40	0.50	14.01	13.00	13.13	True	3.30
56327	0.57	0.56	12.29	12.40	12.52	False	-0.07
56327	0.56	0.56	12.40	12.40	12.52	False	-0.59
56367	0.55	0.61	12.50	11.88	12.01	False	1.29
56367	0.56	0.61	12.40	11.88	12.01	False	0.77
57920	0.48	0.48	13.20	13.20	13.34	False	-0.55
57920	0.51	0.48	12.90	13.20	13.34	False	1.01
75399	0.37	0.38	14.33	14.22	14.37	False	-1.04
75399	0.38	0.38	14.22	14.22	14.37	False	-0.50
75412	0.69	0.69	11.02	11.02	11.13	False	-0.67
75412	0.72	0.69	10.67	11.02	11.13	False	1.13
75424	0.65	0.69	11.46	11.02	11.13	False	0.44
75424	0.67	0.69	11.24	11.02	11.13	False	-0.69
75513	0.72	0.72	10.67	10.67	10.78	False	-0.69
75513	0.70	0.72	10.90	10.67	10.78	False	-0.62
75691	0.88	0.87	8.30	8.49	8.58	False	0.20
75691	0.89	0.87	8.09	8.49	8.58	False	1.27
75693	0.76	0.80	10.17	9.63	9.73	False	1.03
75693	0.77	0.80	10.04	9.63	9.73	False	0.36
75718	0.70	0.68	10.90	11.13	11.24	False	0.51
75718	0.70	0.68	10.90	11.13	11.24	False	0.51
75719	0.57	0.58	12.29	12.19	12.32	False	-1.13
75719	0.53	0.58	12.70	12.19	12.32	False	0.71
75774	0.53	0.51	12.70	12.90	13.03	False	0.47
75774	0.53	0.51	12.70	12.90	13.03	False	0.47
75831	0.83	0.84	9.18	9.02	9.11	False	-0.90
75831	0.83	0.84	9.18	9.02	9.11	False	-0.90
76289	0.79	0.81	9.77	9.49	9.59	False	-0.28
76289	0.77	0.81	10.04	9.49	9.59	False	1.12
76397	0.38	0.41	14.22	13.91	14.05	False	-0.39
76397	0.40	0.41	14.01	13.91	14.05	False	-1.05
76401	0.60	0.61	11.99	11.88	12.01	False	-1.16
76401	0.58	0.61	12.19	11.88	12.01	False	-0.29

Table K-16. 2010–11 MontCAS: Delta Analysis Results – Science Grade 8

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
121209	0.45	0.44	13.50	13.60	1	False	-0.76
121222	0.70	0.68	10.90	11.13	1	False	-0.17
121232	0.59	0.61	12.09	11.88	1	False	-0.46
122023	0.63	0.67	11.67	11.24	1	False	0.14
122036	0.85	0.82	8.85	9.34	1	False	0.75
122736	0.70	0.69	10.90	11.02	1	False	-0.48
122738	0.53	0.49	12.70	13.10	1	False	0.15
144065	0.76	0.74	10.17	10.43	1	False	-0.03
39701	0.71	0.62	10.79	11.78	1	False	1.99
39780	0.85	0.88	8.85	8.30	1	False	0.22
39782	0.55	0.59	12.50	12.09	1	False	0.14
54228	0.89	0.89	8.09	8.09	1	False	-0.54
54228	0.88	0.89	8.30	8.09	1	False	-0.81
54237	0.80	0.74	9.63	10.43	1	False	1.54
54237	0.79	0.74	9.77	10.43	1	False	1.13
56773	0.77	0.79	10.04	9.77	1	False	-0.47
56773	0.75	0.79	10.30	9.77	1	False	0.28
56846	0.65	0.68	11.46	11.13	1	False	-0.17
56846	0.66	0.68	11.35	11.13	1	False	-0.49
56876	0.52	0.57	12.80	12.29	1	False	0.45
56876	0.53	0.57	12.70	12.29	1	False	0.15
56879	0.43	0.44	13.71	13.60	1	False	-0.60
56879	0.44	0.44	13.60	13.60	1	False	-0.89
56894	0.79	0.80	9.77	9.63	1	False	-0.86
56894	0.77	0.80	10.04	9.63	1	False	-0.07
56985	0.45	0.50	13.50	13.00	1	False	0.50
56985	0.47	0.50	13.30	13.00	1	False	-0.08
56989	0.35	0.36	14.54	14.43	1	False	-0.51
56989	0.36	0.36	14.43	14.43	1	False	-0.82
89277	0.93	0.93	7.10	7.10	1	False	-0.45
89361	0.67	0.66	11.24	11.35	1	False	-0.53
89392	0.39	0.37	14.12	14.33	1	False	-0.51
89392	0.40	0.37	14.01	14.33	1	False	-0.21
89406	0.70	0.71	10.90	10.79	1	False	-0.82
89406	0.71	0.71	10.79	10.79	1	False	-0.79
89426	0.64	0.63	11.57	11.67	1	False	-0.57
89426	0.62	0.63	11.78	11.67	1	False	-0.77
89433	0.48	0.56	13.20	12.40	1	False	1.32
89433	0.47	0.56	13.30	12.40	1	False	1.61
89463	0.28	0.34	15.33	14.65	1	False	1.18
89463	0.32	0.34	14.87	14.65	1	False	-0.16
89504	0.69	0.71	11.02	10.79	1	False	-0.49
89504	0.70	0.71	10.90	10.79	1	False	-0.82
89513	0.84	0.82	9.02	9.34	1	False	0.26
89513	0.84	0.82	9.02	9.34	1	False	0.26
89534	0.30	0.23	15.07	15.92	4	False	1.20
89534	0.30	0.23	15.16	15.92	4	False	0.95
89591	0.55	0.57	12.50	12.29	1	False	-0.43
89591	0.55	0.57	12.50	12.29	1	False	-0.43
89621	0.49	0.46	13.10	13.40	1	False	-0.16
89621	0.49	0.46	13.10	13.40	1	False	-0.16
89637	0.51	0.50	12.90	13.00	1	False	-0.71
89637	0.52	0.50	12.80	13.00	1	False	-0.42
89644	0.48	0.52	13.20	12.80	1	False	0.19
89702	0.60	0.72	11.99	10.67	1	False	2.65

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
89702	0.60	0.72	11.99	10.67	1	False	2.65
89710	0.38	0.38	14.22	14.22	1	False	-0.84
89710	0.37	0.38	14.33	14.22	1	False	-0.53
89752	0.73	0.72	10.55	10.67	1	False	-0.43
89766	0.61	0.62	11.88	11.78	1	False	-0.76
89857	0.64	0.65	11.57	11.46	1	False	-0.78
89857	0.66	0.65	11.35	11.46	1	False	-0.54
89859	0.46	0.39	13.40	14.12	1	False	0.97
89859	0.40	0.39	14.01	14.12	1	False	-0.80
89860	0.45	0.29	13.50	15.21	1	True	3.75
89860	0.45	0.29	13.50	15.21	1	True	3.75
89881	0.59	0.58	12.09	12.19	1	False	-0.63
89881	0.57	0.58	12.29	12.19	1	False	-0.73
89892	0.68	0.66	11.13	11.35	1	False	-0.20
89895	0.83	0.82	9.18	9.34	1	False	-0.21
89895	0.82	0.82	9.34	9.34	1	False	-0.66
89900	0.60	0.61	11.99	11.88	1	False	-0.75
89900	0.61	0.61	11.88	11.88	1	False	-0.89
89901	0.41	0.49	13.91	13.10	1	False	1.40
89901	0.44	0.49	13.60	13.10	1	False	0.52
89914	0.41	0.42	13.91	13.81	1	False	-0.58
89914	0.43	0.42	13.71	13.81	1	False	-0.78
89917	0.83	0.84	9.18	9.02	1	False	-0.86
89917	0.82	0.84	9.34	9.02	1	False	-0.41
89934	0.53	0.54	12.70	12.60	1	False	-0.70
89934	0.55	0.54	12.50	12.60	1	False	-0.67
89958	0.39	0.39	14.17	14.09	4	False	-0.62
89958	0.37	0.39	14.35	14.09	4	False	-0.09

Table K-17. 2010–11 MontCAS: Delta Analysis Results – Science Grade 10

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
119656	0.60	0.59	11.99	12.09	1	False	-0.07
119661	0.59	0.59	12.09	12.09	1	False	-0.67
119664	0.38	0.46	14.22	13.40	1	False	2.92
119674	0.56	0.55	12.40	12.50	1	False	-0.02
119675	0.41	0.44	13.91	13.60	1	False	-0.12
119676	0.43	0.39	13.71	14.12	1	False	2.02
119943	0.70	0.69	10.90	11.02	1	False	-0.15
119960	0.35	0.39	14.60	14.17	4	False	0.50
119960	0.33	0.39	14.73	14.17	4	False	1.31
120064	0.78	0.78	9.91	9.91	1	False	-0.97
120311	0.64	0.63	11.57	11.67	1	False	-0.11
130550	0.87	0.88	8.49	8.30	1	False	-0.05
130592	0.60	0.64	11.99	11.57	1	False	0.83
130594	0.57	0.58	12.29	12.19	1	False	-1.12
134479	0.56	0.55	12.40	12.50	1	False	-0.02
134480	0.54	0.51	12.60	12.90	1	False	1.21
134487	0.81	0.84	9.49	9.02	1	False	1.45
134489	0.72	0.72	10.67	10.67	1	False	-0.87
134543	0.60	0.60	11.99	11.99	1	False	-0.69
134545	0.61	0.60	11.88	11.99	1	False	-0.08
134770	0.61	0.61	11.88	11.88	1	False	-0.70
134778	0.41	0.38	13.91	14.22	1	False	1.45

continued

<i>IREF</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old delta</i>	<i>New delta</i>	<i>Maximum</i>	<i>Discard</i>	<i>Standardized difference</i>
134792	0.69	0.72	11.02	10.67	1	False	0.53
134799	0.77	0.77	10.04	10.04	1	False	-0.95
52953	0.30	0.32	15.13	14.93	4	False	-0.93
52953	0.33	0.32	14.82	14.93	4	False	0.37
52960	0.42	0.44	13.81	13.60	1	False	-0.72
52960	0.43	0.44	13.71	13.60	1	False	-1.06
75433	0.52	0.54	12.80	12.60	1	False	-0.60
75441	0.54	0.56	12.60	12.40	1	False	-0.57
75441	0.54	0.56	12.60	12.40	1	False	-0.57
75452	0.49	0.51	13.10	12.90	1	False	-0.64
75452	0.48	0.51	13.20	12.90	1	False	-0.06
75620	0.36	0.38	14.43	14.22	1	False	-0.76
75630	0.45	0.47	13.50	13.30	1	False	-0.69
75630	0.42	0.47	13.81	13.30	1	False	1.10
75632	0.47	0.51	13.30	12.90	1	False	0.53
75632	0.50	0.51	13.00	12.90	1	False	-1.15
75634	0.44	0.46	13.60	13.40	1	False	-0.70
75635	0.52	0.53	12.80	12.70	1	False	-1.18
75644	0.40	0.42	14.01	13.81	1	False	-0.74
75644	0.36	0.42	14.43	13.81	1	False	1.73
75725	0.36	0.36	14.43	14.43	1	False	-0.35
75725	0.34	0.36	14.65	14.43	1	False	-0.76
75726	0.45	0.40	13.50	14.01	1	False	2.59
75726	0.46	0.40	13.40	14.01	1	True	3.18
75777	0.75	0.77	10.30	10.04	1	False	0.08
75777	0.78	0.77	9.91	10.04	1	False	-0.17
75786	0.70	0.73	10.90	10.55	1	False	0.58
75786	0.69	0.73	11.02	10.55	1	False	1.25
75787	0.82	0.80	9.34	9.63	1	False	0.72
75787	0.82	0.80	9.34	9.63	1	False	0.72
75803	0.55	0.57	12.50	12.29	1	False	-0.55
75803	0.58	0.57	12.19	12.29	1	False	-0.05
75804	0.47	0.49	13.30	13.10	1	False	-0.67
75846	0.72	0.72	10.67	10.67	1	False	-0.87
75846	0.71	0.72	10.79	10.67	1	False	-0.82
75854	0.63	0.66	11.67	11.35	1	False	0.28
75859	0.58	0.60	12.19	11.99	1	False	-0.49
75863	0.64	0.65	11.57	11.46	1	False	-0.99
75865	0.51	0.51	12.90	12.90	1	False	-0.56
75865	0.50	0.51	13.00	12.90	1	False	-1.15
75875	0.80	0.81	9.63	9.49	1	False	-0.50
75875	0.83	0.81	9.18	9.49	1	False	0.76
75941	0.71	0.70	10.79	10.90	1	False	-0.16
75963	0.60	0.58	11.99	12.19	1	False	0.55
75970	0.54	0.53	12.60	12.70	1	False	0.00
75976	0.77	0.78	10.04	9.91	1	False	-0.63
75976	0.78	0.78	9.91	9.91	1	False	-0.97

Table K-18. 2010–11 MontCAS: Rescore Analysis Results by Subject and Grade

<i>Subject</i>	<i>Grade</i>	<i>IREF</i>	<i>Maximum</i>	<i>Old mean</i>	<i>New mean</i>	<i>Old standard deviation</i>	<i>New standard deviation</i>	<i>Effect size</i>	<i>Discard</i>
Mathematics	3	138799	4	1.80	1.67	1.32	1.29	-0.10	No
	4	140183	4	2.18	2.00	1.30	1.28	-0.14	No
	6	63054	4	1.45	1.44	1.19	1.19	-0.01	No
	8	87874	4	2.20	2.18	1.22	1.24	-0.02	No
	10	62391	4	0.87	0.91	1.00	0.98	0.04	No
		59407	4	1.27	1.22	1.26	1.25	-0.04	No
Reading	3	151240	4	1.26	1.29	0.99	0.93	0.03	No
		150991	4	1.72	1.46	1.04	0.92	-0.24	No
	4	151772	4	1.63	1.76	0.84	0.65	0.15	No
		151659	4	1.39	1.49	0.93	0.74	0.10	No
	5	150456	4	1.80	1.84	0.96	0.81	0.04	No
		150518	4	2.01	2.05	0.95	1.01	0.04	No
	6	95469	4	1.82	1.82	0.92	0.89	-0.01	No
		68294	4	1.37	1.55	0.98	0.87	0.18	No
	7	148887	4	1.96	2.09	1.18	0.97	0.11	No
		149016	4	1.99	2.22	1.11	1.03	0.21	No
	8	149368	4	2.10	2.24	1.06	1.16	0.13	No
		149193	4	2.28	2.42	1.05	1.06	0.13	No
	10	149566	4	1.81	1.96	1.12	1.04	0.14	No
		149648	4	2.00	2.11	1.21	1.07	0.09	No
Science	4	120312	4	1.99	1.89	0.97	0.88	-0.11	No
	8	89534	4	1.23	0.99	1.06	0.94	-0.24	No
		89958	4	1.50	1.45	1.21	1.22	-0.04	No
	10	119960	4	1.46	1.50	1.15	1.14	0.04	No
		52953	4	1.38	1.11	0.91	0.75	-0.30	No

Appendix L—SCORE DISTRIBUTIONS

Table L-1. 2010–11 MontCAS: Performance Level Distributions by Subject and Grade

<i>Subject</i>	<i>Grade</i>	<i>Performance level</i>	<i>Percent in level:</i>		
			<i>2010–10</i>	<i>2008–09</i>	<i>2007–08</i>
Mathematics	3	4	29.82	28.91	27.81
		3	40.43	40.49	39.66
		2	15.93	16.15	16.04
		1	13.82	14.44	16.49
	4	4	31.37	31.97	28.06
		3	38.79	37.30	39.10
		2	16.55	17.00	18.04
		1	13.30	13.73	14.80
	5	4	34.77	32.57	32.98
		3	37.63	37.93	34.39
		2	15.96	16.50	18.47
		1	11.64	13.00	14.16
	6	4	33.12	32.42	33.08
		3	33.64	36.24	32.04
		2	20.27	17.76	18.68
		1	12.96	13.59	16.20
	7	4	36.89	35.92	31.30
		3	32.83	31.69	35.40
		2	16.29	16.39	18.72
		1	14.00	16.00	14.58
	8	4	27.06	26.88	27.43
		3	38.85	40.14	33.53
		2	22.35	22.94	23.81
		1	11.74	10.05	15.22
	10	4	25.20	21.67	18.93
		3	34.19	35.74	35.95
		2	29.83	31.27	35.30
		1	10.78	11.32	9.82
Reading	3	4	46.28	45.60	40.98
		3	39.30	39.45	43.78
		2	10.62	12.48	12.23
		1	3.79	2.47	3.00
	4	4	43.95	45.29	39.43
		3	39.32	38.21	42.21
		2	12.24	12.16	13.96
		1	4.49	4.35	4.40
	5	4	57.23	55.15	48.45
		3	30.20	31.32	35.96
		2	8.57	9.88	11.64
		1	4.00	3.64	3.95
	6	4	53.17	53.96	49.91
		3	34.44	32.68	35.23
		2	8.32	8.65	9.15
		1	4.08	4.71	5.70
	7	4	55.36	48.61	47.15
		3	30.64	35.75	36.41
		2	8.87	10.31	11.73
		1	5.13	5.33	4.71

Subject	Grade	Performance level	Percent in level:		
			2010–10	2008–09	2007–08
Reading	8	4	55.77	54.42	48.56
		3	28.71	30.23	33.05
		2	8.69	9.54	11.27
		1	6.83	5.80	7.12
	10	4	49.18	44.01	43.77
		3	33.61	36.82	35.19
		2	9.33	9.68	11.67
		1	7.88	9.49	9.37
Science	4	4	13.08	15.83	11.62
		3	48.86	50.71	54.49
		2	30.31	27.73	27.01
		1	7.75	5.73	6.88
	8	4	14.78	18.41	14.16
		3	49.73	44.36	46.13
		2	24.01	27.71	28.43
		1	11.49	9.51	11.29
	10	4	20.17	17.82	18.27
		3	27.15	25.02	24.41
		2	33.06	33.66	33.73
		1	19.62	23.50	23.59

Figure L-1. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 3

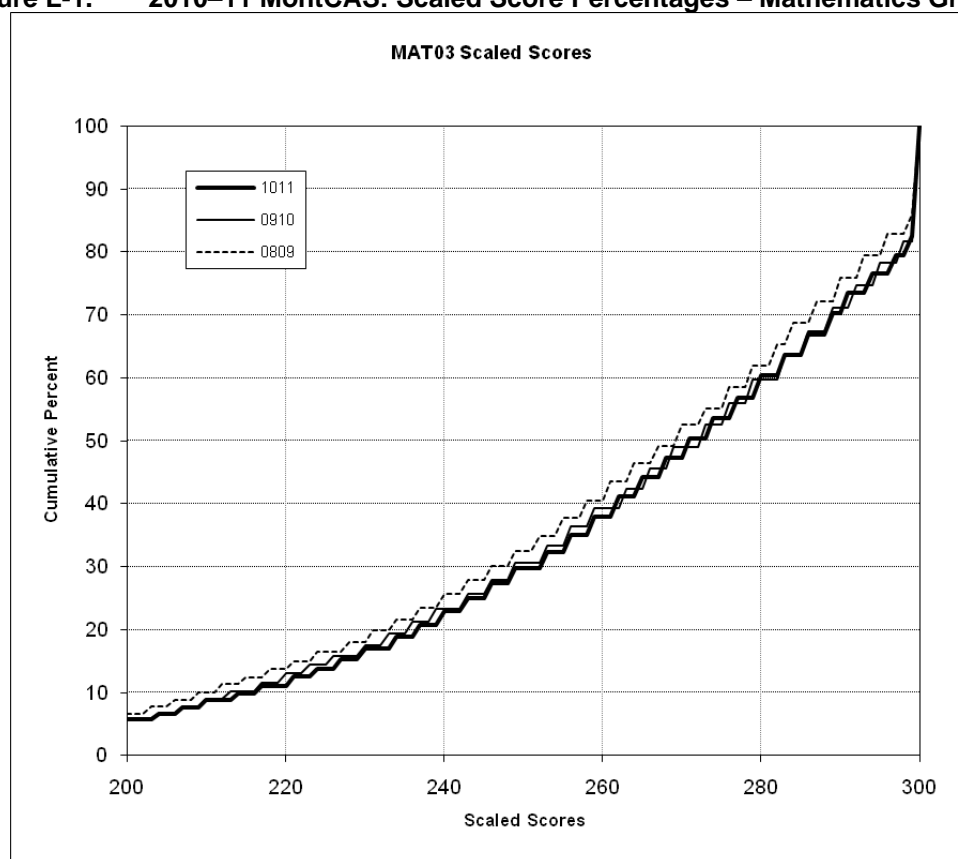


Figure L-2. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 4

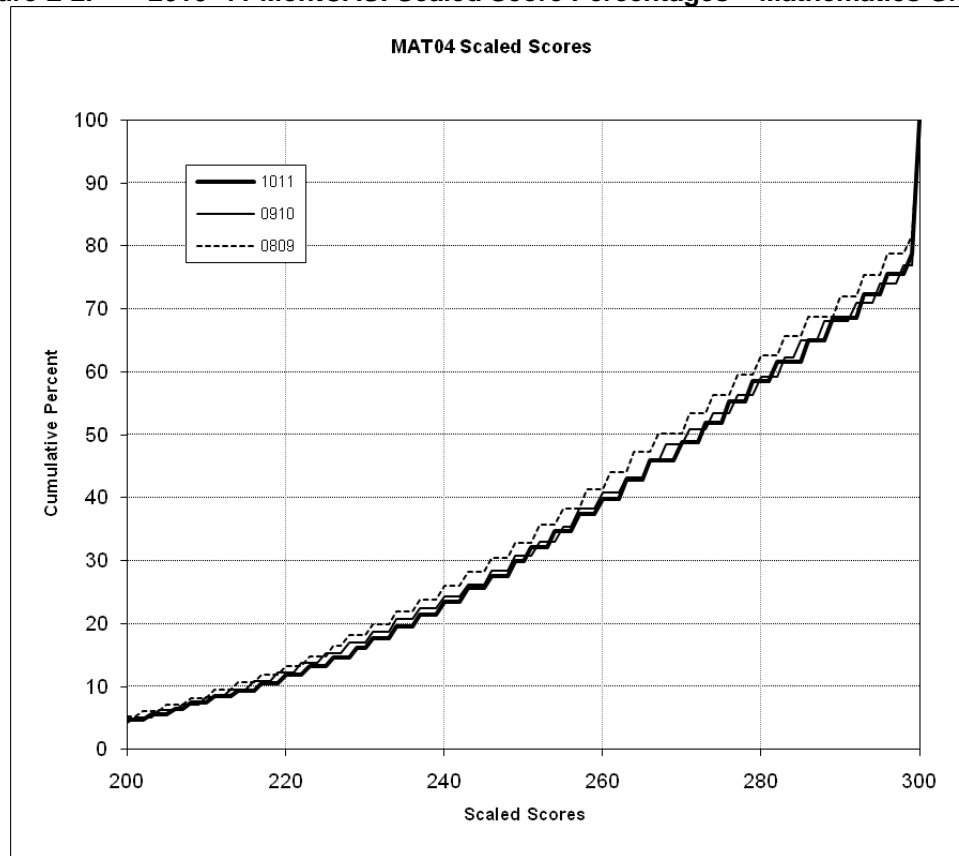


Figure L-3. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 5

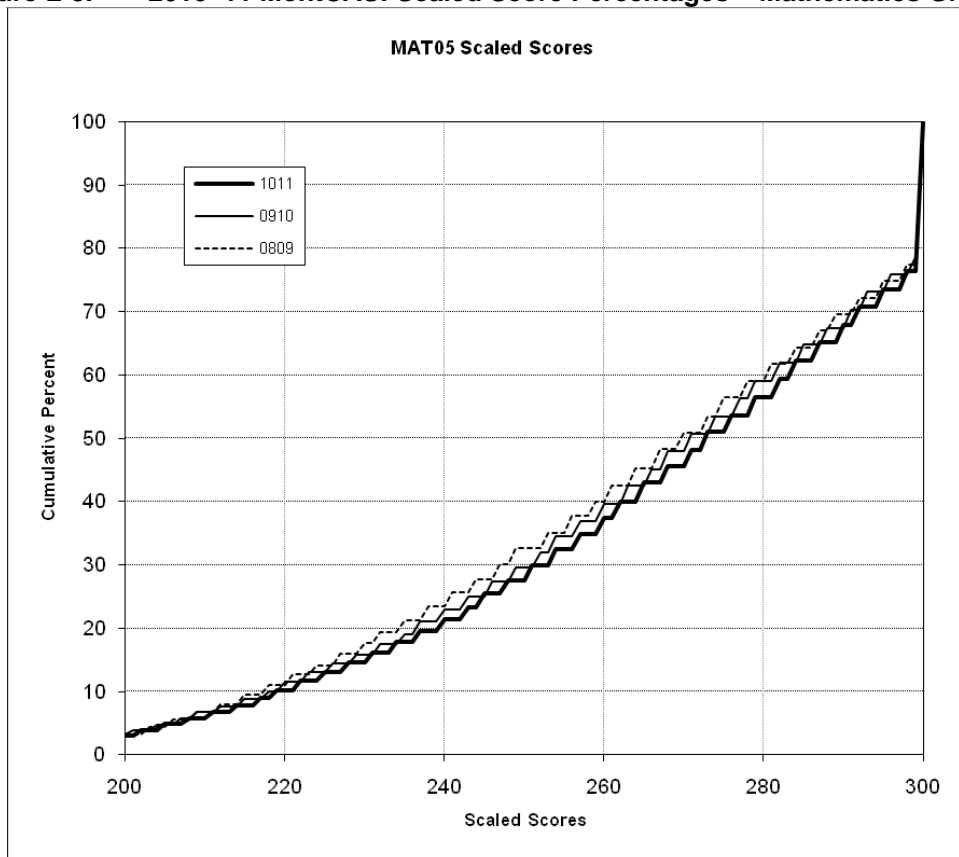


Figure L-4. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 6

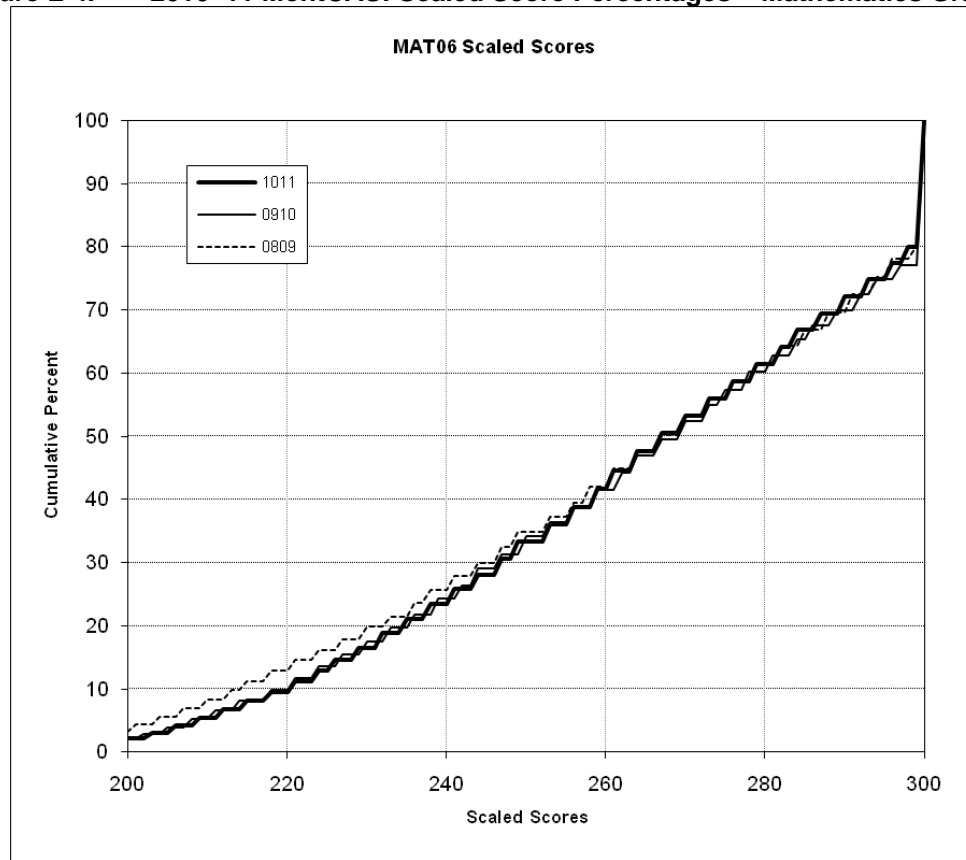


Figure L-5. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 7

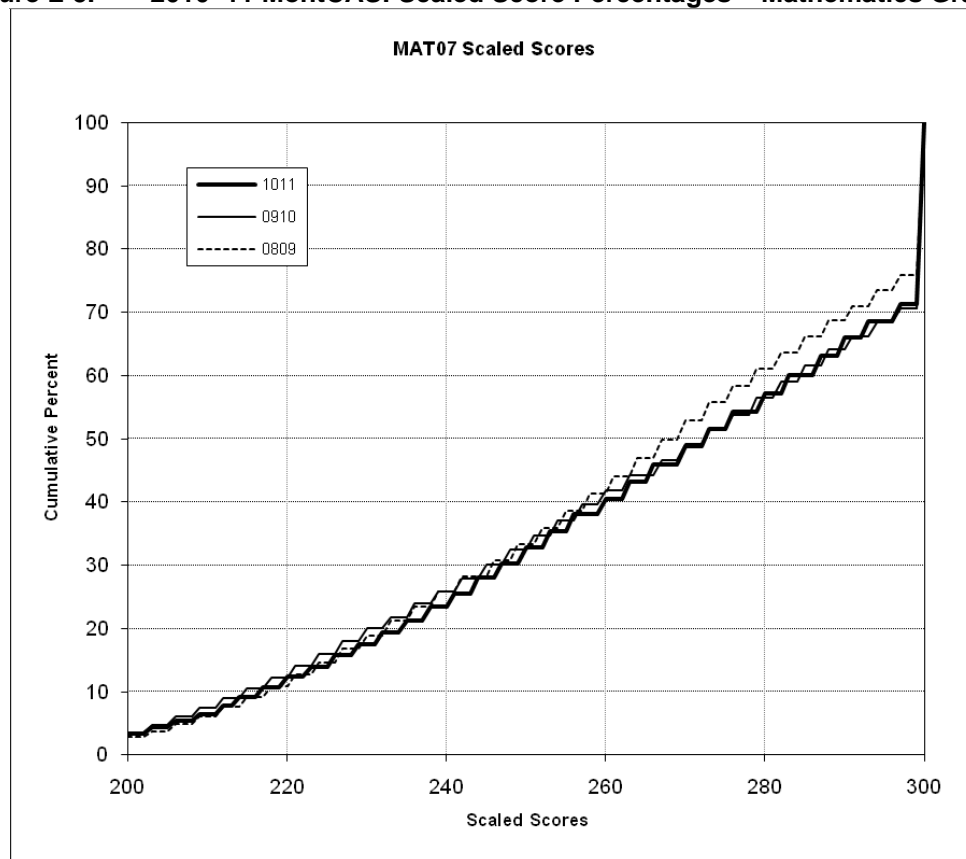


Figure L-6. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 8

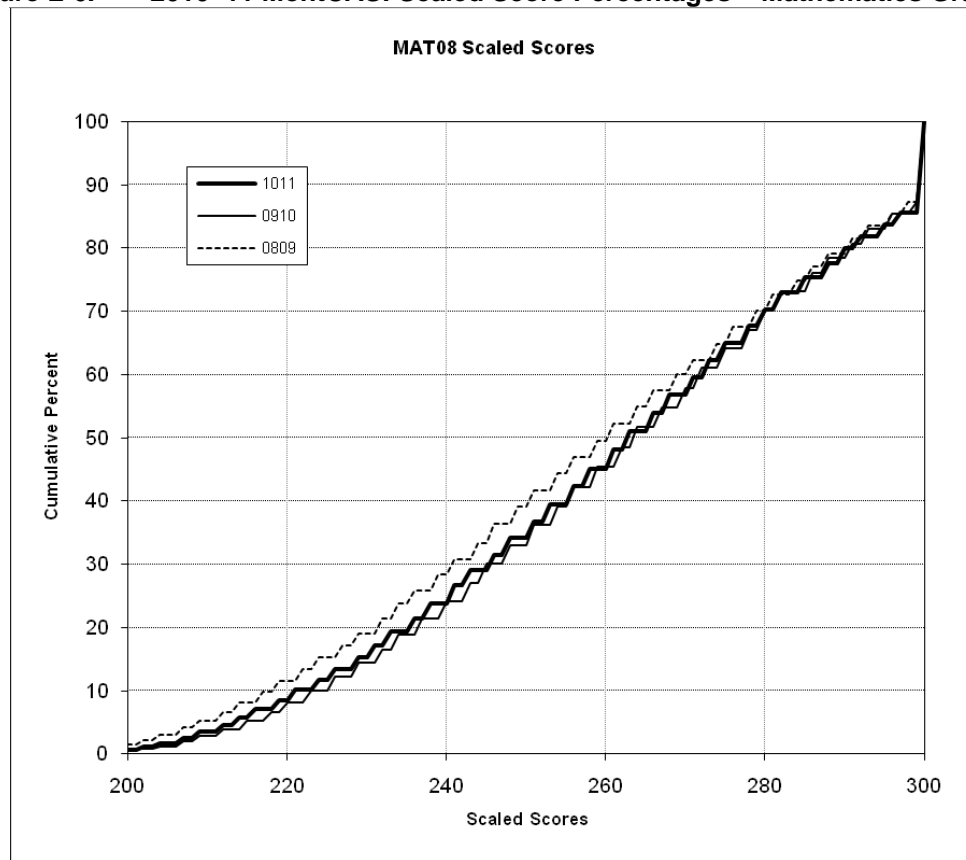


Figure L-7. 2010–11 MontCAS: Scaled Score Percentages – Mathematics Grade 10

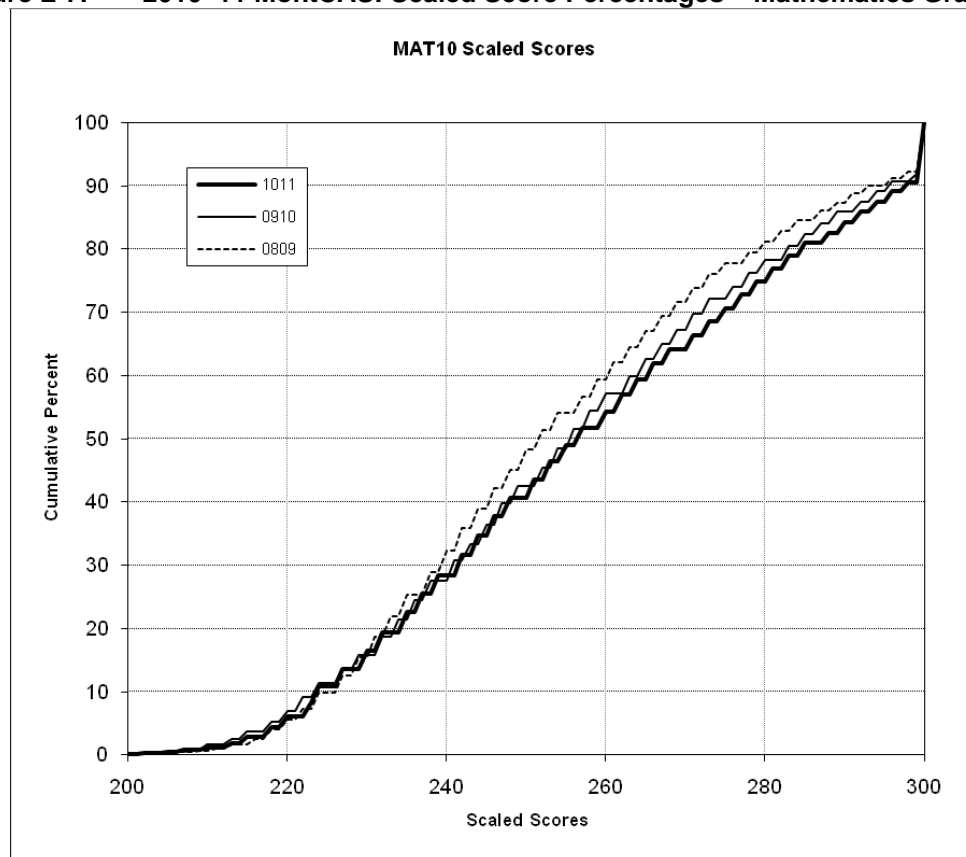


Figure L-8. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 3

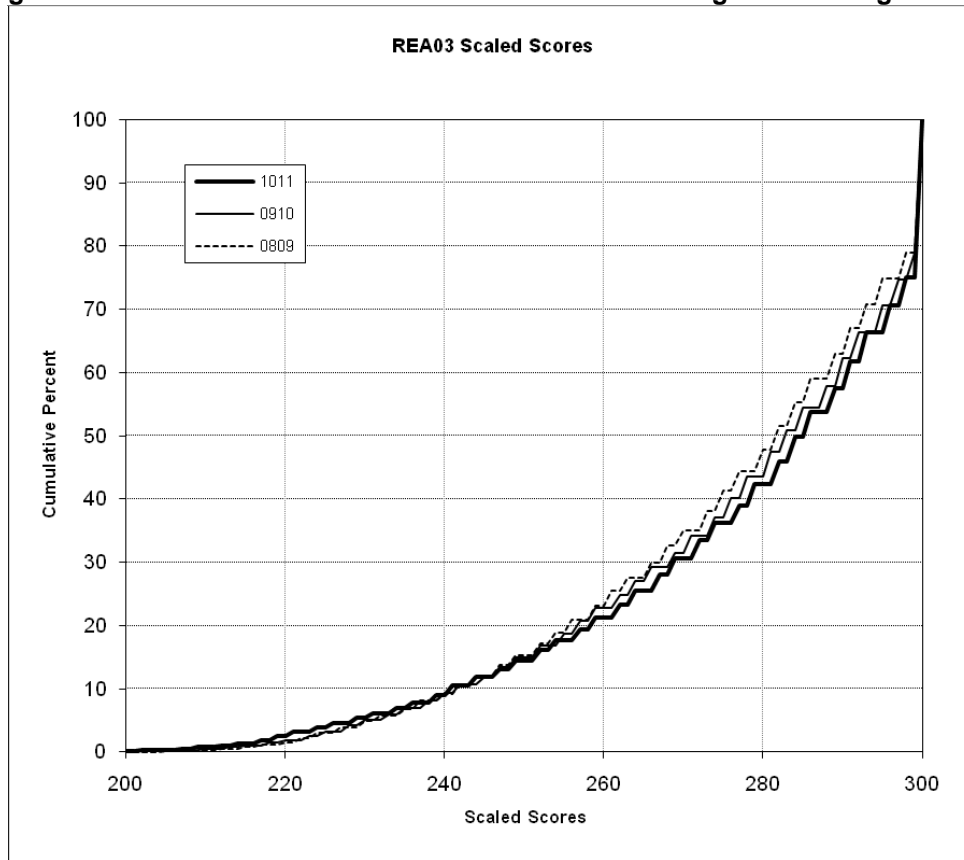


Figure L-9. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 4

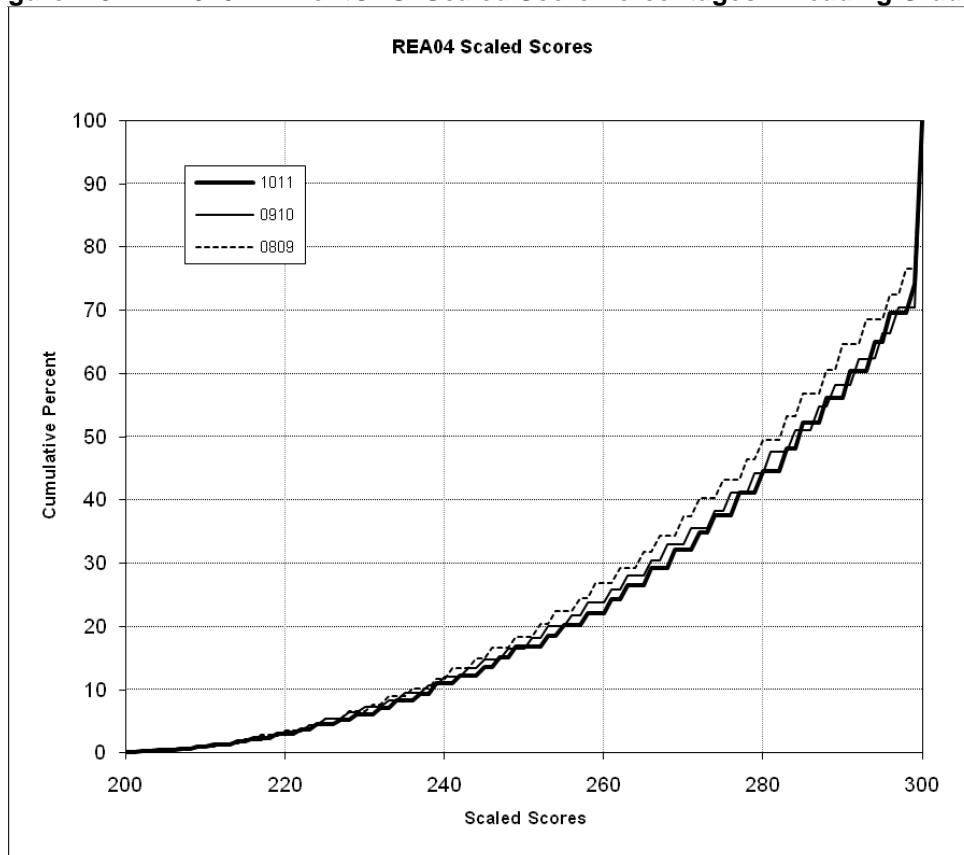


Figure L-10. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 5

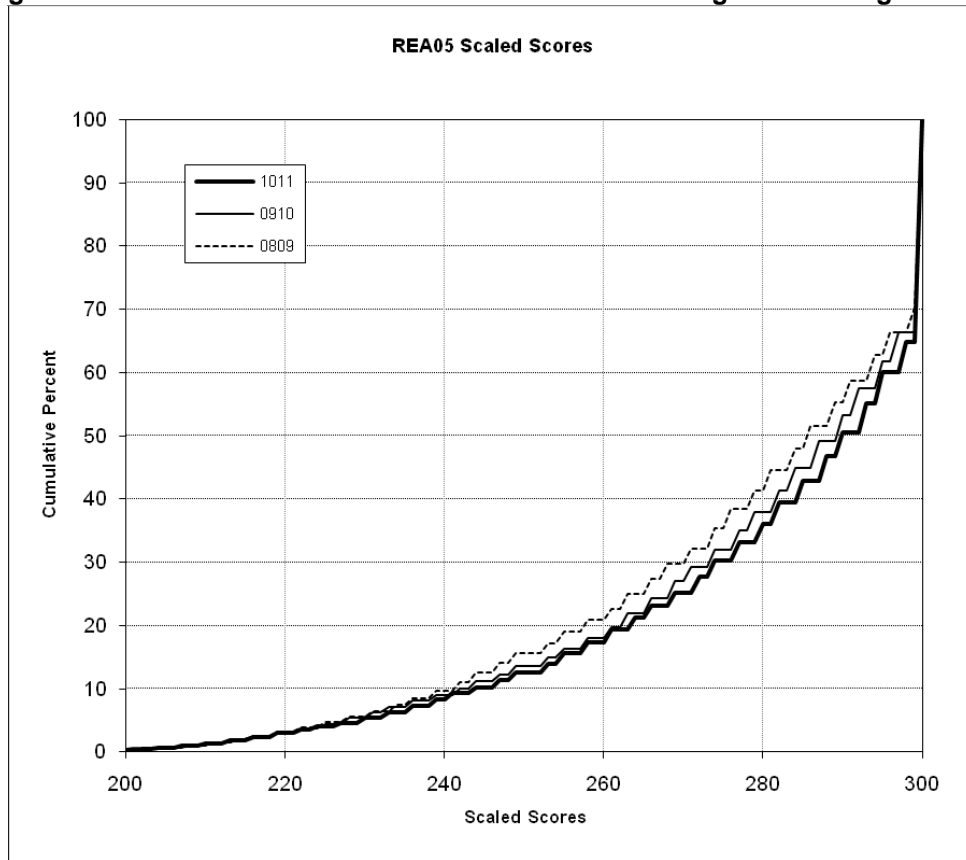


Figure L-11. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 6

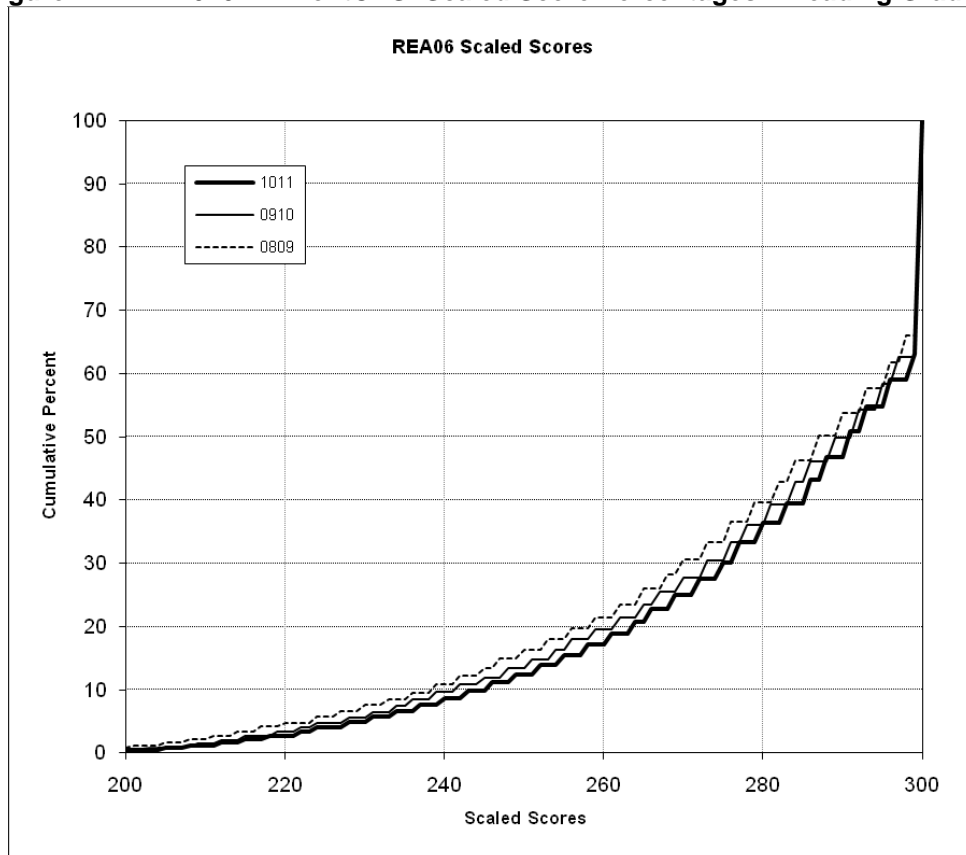


Figure L-12. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 7

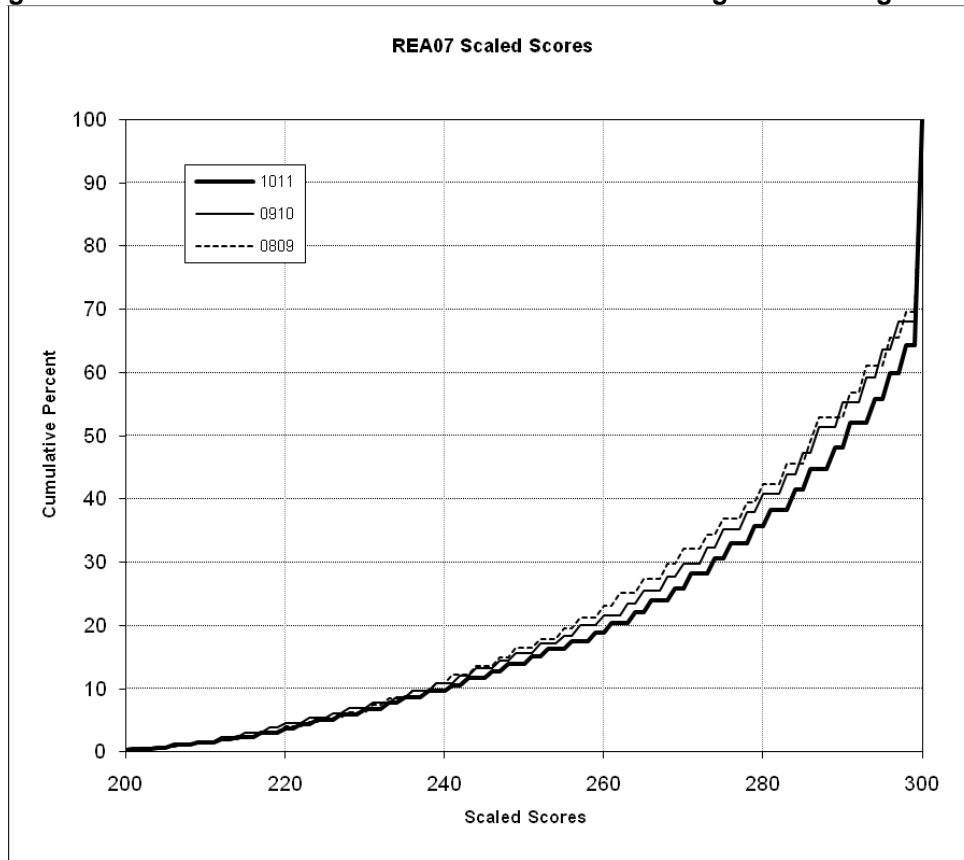


Figure L-13. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 8

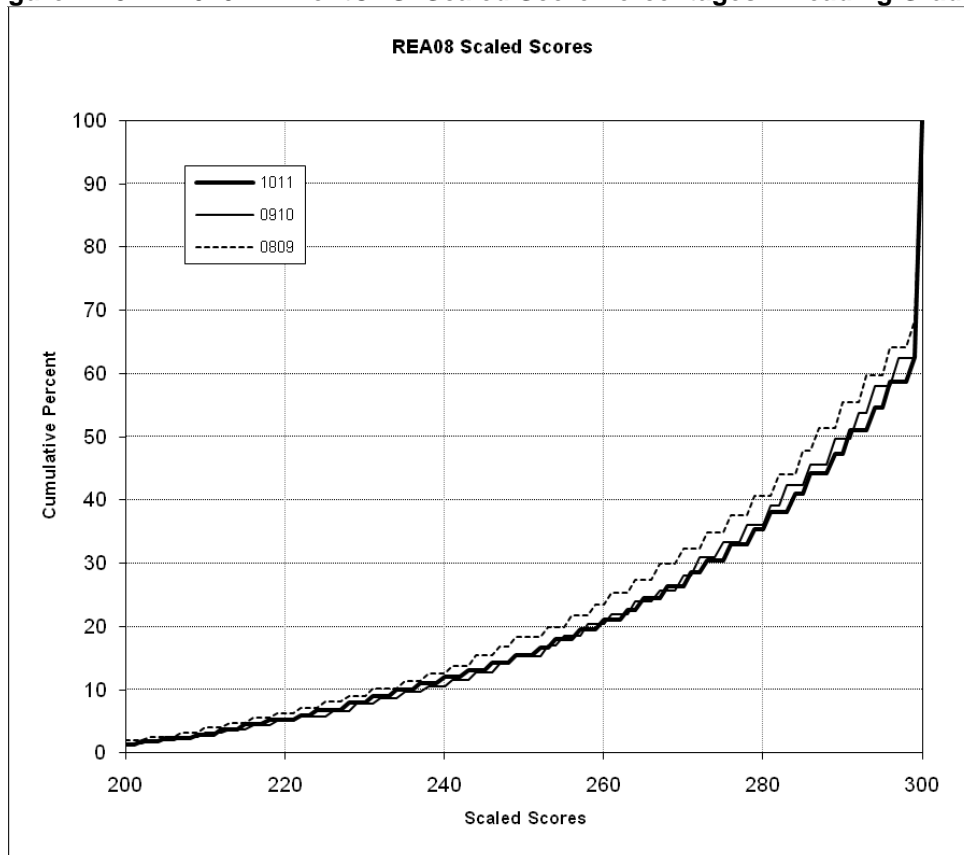


Figure L-14. 2010–11 MontCAS: Scaled Score Percentages – Reading Grade 10

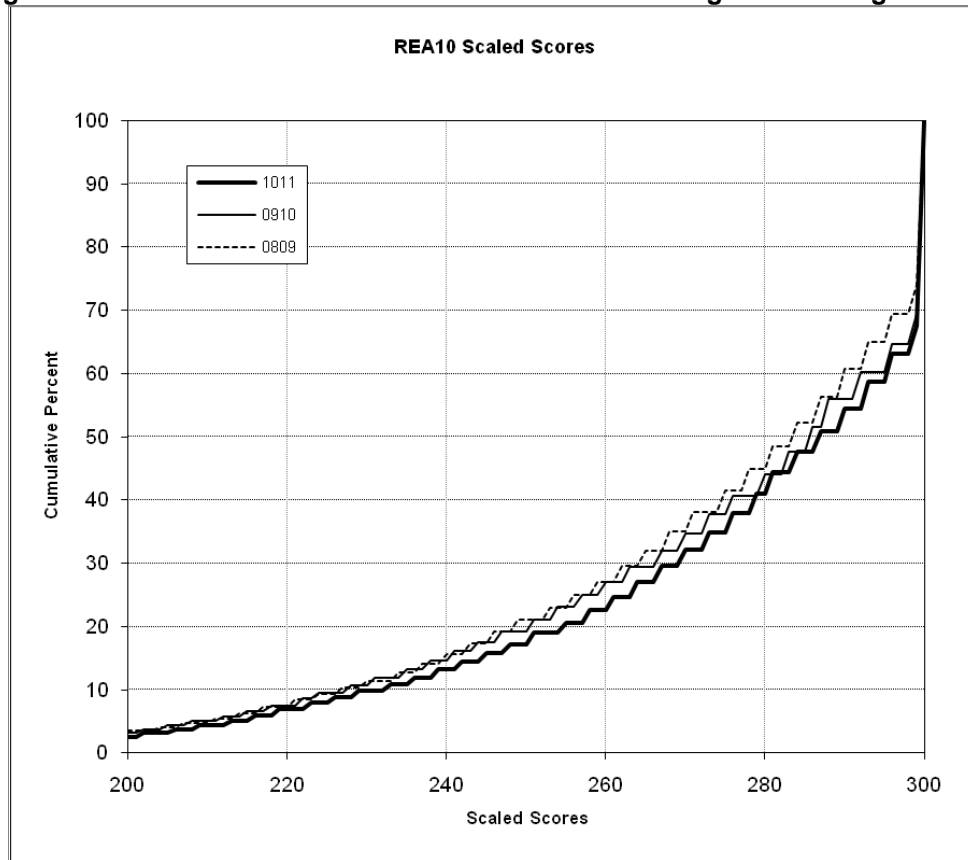


Figure L-15. 2010–11 MontCAS: Scaled Score Percentages – Science Grade 4

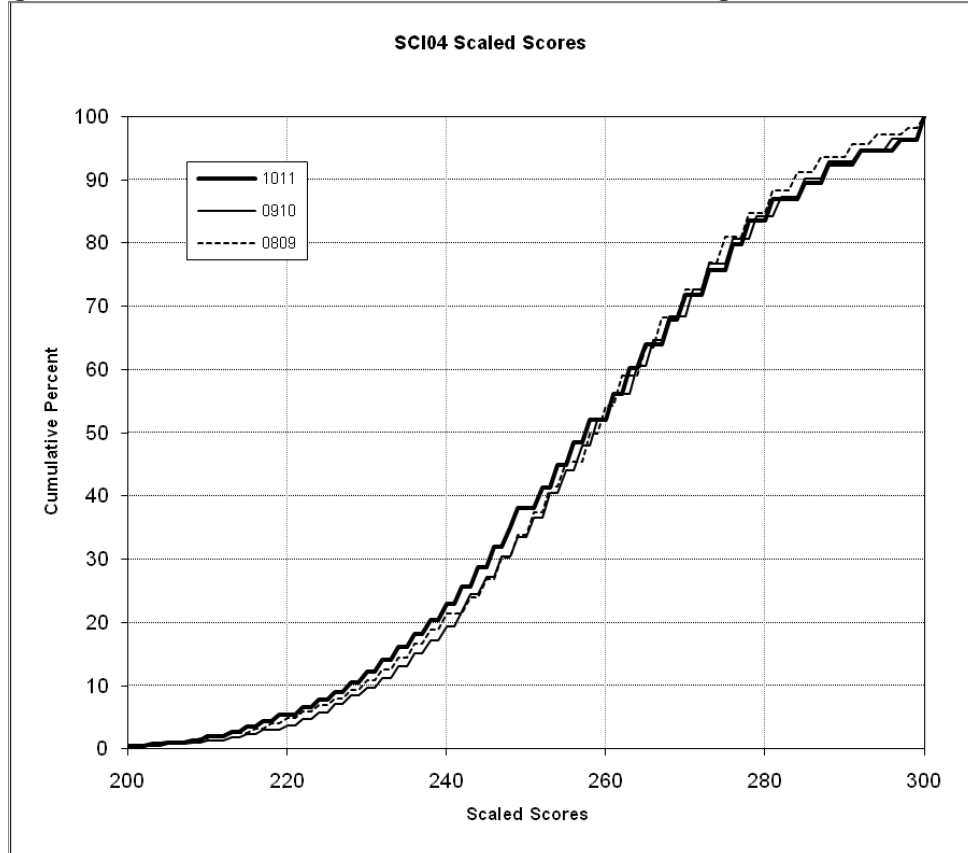


Figure L-16. 2010–11 MontCAS: Scaled Score Percentages – Science Grade 8

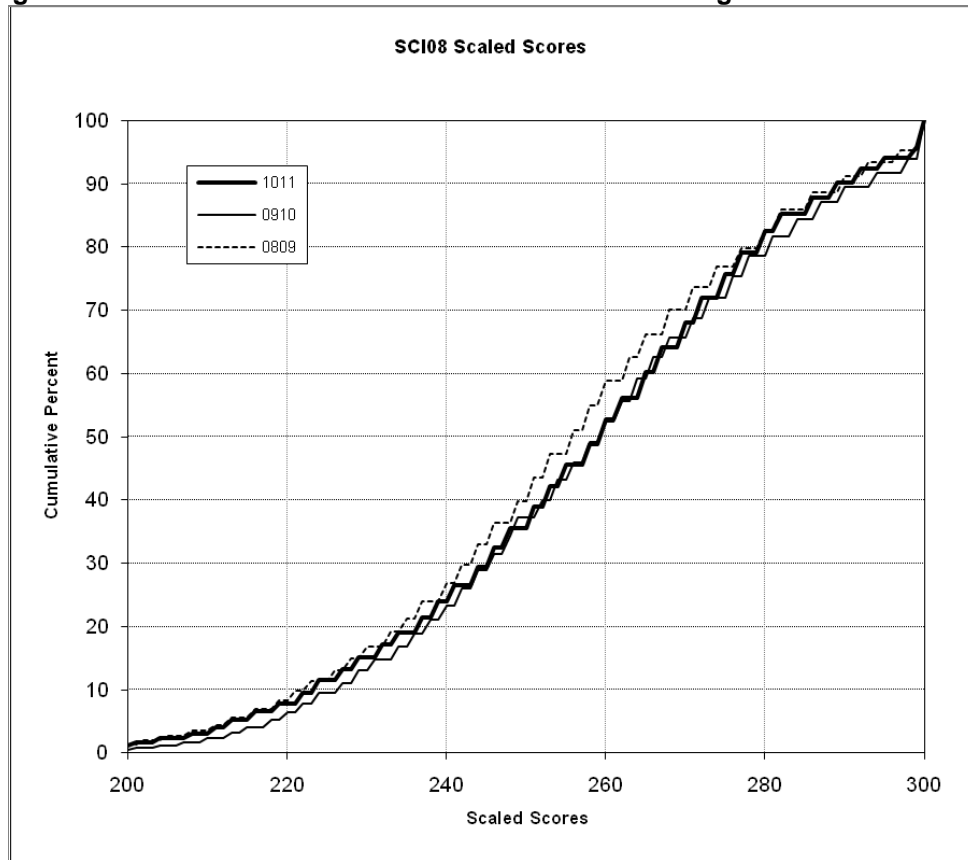
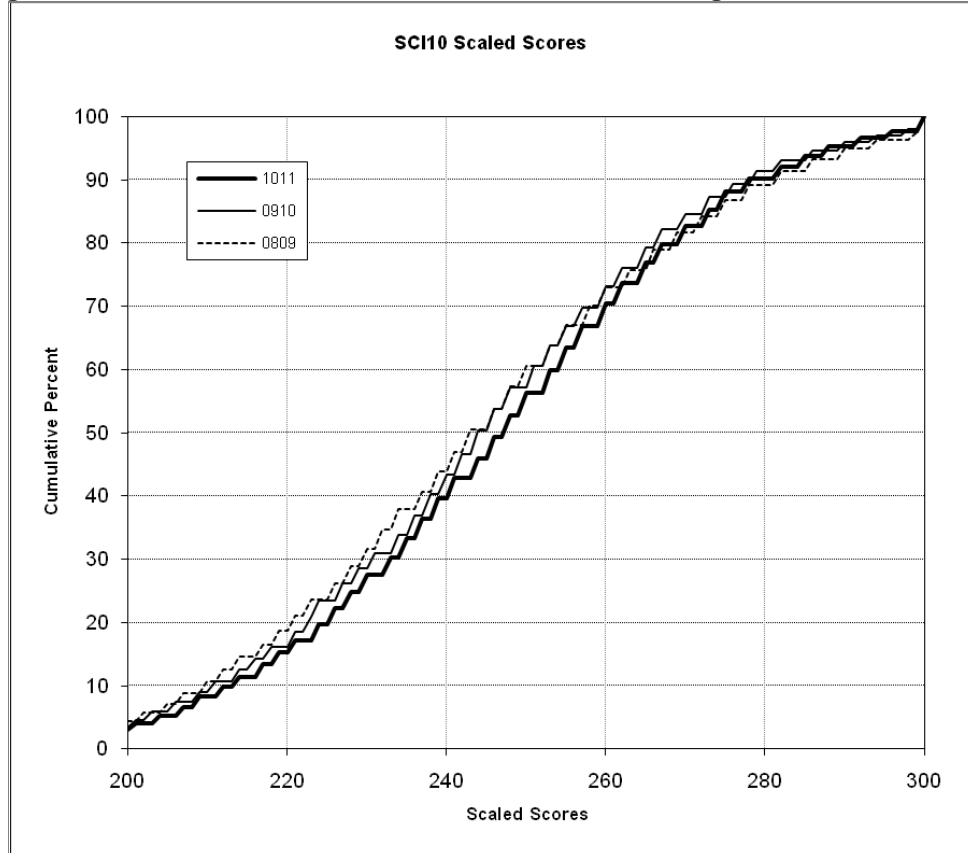


Figure L-17. 2010–11 MontCAS: Scaled Score Percentages – Science Grade 10



Appendix M—RAW TO SCALED SCORE LOOK-UP TABLES

Table M-1. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 3

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	200	1	200	210	200	1	200	210
16	200	1	200	210	200	1	200	210
17	200	1	200	210	200	1	200	210
18	200	1	200	210	200	1	200	210
19	200	1	200	210	200	1	200	210
20	200	1	200	210	200	1	200	210
21	200	1	200	210	200	1	200	210
22	200	1	200	210	200	1	200	210
23	200	1	200	210	200	1	200	210
24	204	1	200	214	200	1	200	210
25	207	1	200	217	200	1	200	210
26	210	1	200	220	204	1	200	214
27	214	1	204	224	207	1	200	217
28	217	1	207	227	210	1	200	220
29	221	1	211	231	213	1	203	223
30	224	1	214	234	217	1	207	227
31	227	2	217	237	220	1	210	230
32	230	2	220	240	223	1	213	233
33	234	2	224	244	226	2	216	236
34	237	2	227	247	230	2	220	240
35	240	2	230	250	233	2	223	243
36	243	2	233	253	236	2	226	246
37	246	2	236	256	239	2	229	249
38	249	2	239	259	243	2	233	253
39	253	3	243	263	246	2	236	256
40	256	3	246	266	249	2	239	259
41	259	3	249	269	253	3	243	263
42	262	3	252	272	256	3	246	266
43	265	3	255	275	259	3	249	269
44	268	3	258	278	263	3	253	273
45	271	3	261	281	266	3	256	276
46	274	3	264	284	269	3	259	279
47	277	3	267	287	273	3	263	283
48	280	3	270	290	276	3	266	286
49	283	3	273	293	279	3	269	289
50	286	3	276	296	283	3	273	293
51	289	3	279	299	286	3	276	296
52	291	4	281	300	289	3	279	299
53	294	4	284	300	292	4	282	300
54	297	4	287	300	295	4	285	300

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
55	299	4	289	300	298	4	288	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-2. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 4

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	200	1	200	210	200	1	200	210
16	200	1	200	210	200	1	200	210
17	200	1	200	210	200	1	200	210
18	200	1	200	210	200	1	200	210
19	200	1	200	210	200	1	200	210
20	200	1	200	210	201	1	200	211
21	203	1	200	213	204	1	200	214
22	206	1	200	216	207	1	200	217
23	208	1	200	218	210	1	200	220
24	211	1	201	221	213	1	203	223
25	214	1	204	224	216	1	206	226
26	217	1	207	227	219	1	209	229
27	220	1	210	230	222	1	212	232
28	223	1	213	233	225	2	215	235
29	226	2	216	236	228	2	218	238
30	229	2	219	239	231	2	221	241
31	231	2	221	241	234	2	224	244
32	234	2	224	244	237	2	227	247
33	237	2	227	247	240	2	230	250
34	240	2	230	250	243	2	233	253
35	243	2	233	253	246	2	236	256
36	246	2	236	256	249	2	239	259
37	249	2	239	259	252	3	242	262

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
38	251	3	241	261	255	3	245	265
39	254	3	244	264	257	3	247	267
40	257	3	247	267	260	3	250	270
41	260	3	250	270	263	3	253	273
42	263	3	253	273	266	3	256	276
43	266	3	256	276	268	3	258	278
44	270	3	260	280	271	3	261	281
45	273	3	263	283	274	3	264	284
46	276	3	266	286	277	3	267	287
47	279	3	269	289	280	3	270	290
48	282	3	272	292	283	3	273	293
49	286	3	276	296	285	3	275	295
50	289	3	279	299	288	3	278	298
51	293	4	283	300	292	4	282	300
52	296	4	286	300	295	4	285	300
53	299	4	289	300	298	4	288	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-3. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 5

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	200	1	200	210	200	1	200	210
16	200	1	200	210	200	1	200	210
17	202	1	200	212	201	1	200	211
18	205	1	200	215	204	1	200	214
19	208	1	200	218	207	1	200	217
20	211	1	201	221	209	1	200	219

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
21	214	1	204	224	212	1	202	222
22	217	1	207	227	215	1	205	225
23	219	1	209	229	218	1	208	228
24	222	1	212	232	220	1	210	230
25	225	2	215	235	223	1	213	233
26	228	2	218	238	226	2	216	236
27	231	2	221	241	229	2	219	239
28	234	2	224	244	232	2	222	242
29	237	2	227	247	235	2	225	245
30	240	2	230	250	237	2	227	247
31	243	2	233	253	240	2	230	250
32	245	2	235	255	243	2	233	253
33	248	2	238	258	246	2	236	256
34	251	3	241	261	249	2	239	259
35	254	3	244	264	252	3	242	262
36	257	3	247	267	254	3	244	264
37	260	3	250	270	257	3	247	267
38	262	3	252	272	260	3	250	270
39	265	3	255	275	263	3	253	273
40	268	3	258	278	266	3	256	276
41	271	3	261	281	268	3	258	278
42	273	3	263	283	271	3	261	281
43	276	3	266	286	274	3	264	284
44	279	3	269	289	277	3	267	287
45	282	3	272	292	279	3	269	289
46	284	3	274	294	282	3	272	292
47	287	3	277	297	285	3	275	295
48	290	4	280	300	288	3	278	298
49	292	4	282	300	291	4	281	300
50	295	4	285	300	293	4	283	300
51	298	4	288	300	296	4	286	300
52	300	4	290	300	299	4	289	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-4. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 6

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	203	1	200	213	202	1	200	212
16	206	1	200	216	205	1	200	215
17	209	1	200	219	208	1	200	218
18	212	1	202	222	211	1	201	221
19	215	1	205	225	214	1	204	224
20	218	1	208	228	218	1	208	228
21	221	1	211	231	221	1	211	231
22	224	1	214	234	224	1	214	234
23	226	2	216	236	227	2	217	237
24	229	2	219	239	230	2	220	240
25	232	2	222	242	233	2	223	243
26	235	2	225	245	236	2	226	246
27	238	2	228	248	239	2	229	249
28	241	2	231	251	242	2	232	252
29	244	2	234	254	244	2	234	254
30	247	2	237	257	247	2	237	257
31	249	2	239	259	250	3	240	260
32	253	3	243	263	253	3	243	263
33	256	3	246	266	256	3	246	266
34	259	3	249	269	259	3	249	269
35	261	3	251	271	262	3	252	272
36	264	3	254	274	264	3	254	274
37	267	3	257	277	267	3	257	277
38	270	3	260	280	270	3	260	280
39	273	3	263	283	273	3	263	283
40	276	3	266	286	275	3	265	285
41	279	3	269	289	278	3	268	288
42	282	3	272	292	281	3	271	291
43	284	3	274	294	284	3	274	294
44	287	4	277	297	286	3	276	296
45	290	4	280	300	289	4	279	299
46	293	4	283	300	292	4	282	300
47	296	4	286	300	294	4	284	300
48	298	4	288	300	297	4	287	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-5. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 7

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	200	1	200	210	203	1	200	213
16	203	1	200	213	206	1	200	216
17	206	1	200	216	209	1	200	219
18	209	1	200	219	212	1	202	222
19	212	1	202	222	215	1	205	225
20	214	1	204	224	218	1	208	228
21	217	1	207	227	221	1	211	231
22	220	1	210	230	224	1	214	234
23	223	1	213	233	227	2	217	237
24	226	2	216	236	230	2	220	240
25	229	2	219	239	233	2	223	243
26	232	2	222	242	236	2	226	246
27	235	2	225	245	239	2	229	249
28	238	2	228	248	242	2	232	252
29	241	2	231	251	245	2	235	255
30	244	2	234	254	248	2	238	258
31	247	2	237	257	251	3	241	261
32	250	3	240	260	254	3	244	264
33	253	3	243	263	257	3	247	267
34	256	3	246	266	260	3	250	270
35	260	3	250	270	263	3	253	273
36	263	3	253	273	267	3	257	277
37	266	3	256	276	270	3	260	280
38	270	3	260	280	273	3	263	283
39	273	3	263	283	276	3	266	286
40	276	3	266	286	279	3	269	289
41	280	3	270	290	282	3	272	292
42	283	3	273	293	285	3	275	295

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
43	287	3	277	297	288	3	278	298
44	290	4	280	300	291	4	281	300
45	293	4	283	300	294	4	284	300
46	297	4	287	300	297	4	287	300
47	300	4	290	300	300	4	290	300
48	300	4	290	300	300	4	290	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-6. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 8

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	201	1	200	211
12	202	1	200	212	204	1	200	214
13	204	1	200	214	207	1	200	217
14	207	1	200	217	209	1	200	219
15	209	1	200	219	212	1	202	222
16	212	1	202	222	215	1	205	225
17	214	1	204	224	218	1	208	228
18	216	1	206	226	220	1	210	230
19	219	1	209	229	223	1	213	233
20	221	1	211	231	226	2	216	236
21	224	1	214	234	229	2	219	239
22	226	2	216	236	232	2	222	242
23	229	2	220	238	234	2	224	244
24	231	2	222	240	237	2	227	247
25	233	2	224	242	240	2	230	250

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
26	236	2	227	245	243	2	233	253
27	238	2	229	247	245	2	235	255
28	241	2	232	250	248	2	238	258
29	243	2	234	252	251	3	241	261
30	246	2	237	255	254	3	245	263
31	248	2	239	257	256	3	247	265
32	251	3	242	260	259	3	250	268
33	253	3	244	262	262	3	253	271
34	256	3	247	265	264	3	255	273
35	258	3	249	267	267	3	257	277
36	261	3	252	270	270	3	260	280
37	263	3	254	272	272	3	262	282
38	266	3	257	275	275	3	265	285
39	268	3	259	277	278	3	268	288
40	271	3	262	280	280	3	270	290
41	273	3	264	282	282	3	272	292
42	275	3	266	284	286	4	276	296
43	278	3	269	287	288	4	278	298
44	280	3	271	289	291	4	281	300
45	282	3	273	291	293	4	283	300
46	285	4	275	295	296	4	286	300
47	288	4	278	298	299	4	289	300
48	290	4	280	300	300	4	290	300
49	292	4	282	300	300	4	290	300
50	295	4	285	300	300	4	290	300
51	297	4	287	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300				
63	300	4	290	300				
64	300	4	290	300				
65	300	4	290	300				
66	300	4	290	300				

Table M-7. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 10

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	202	1	200	212	200	1	200	210
8	205	1	200	215	203	1	200	213

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
9	207	1	200	217	205	1	200	215
10	210	1	200	220	208	1	200	218
11	213	1	203	223	210	1	200	220
12	215	1	205	225	213	1	203	223
13	218	1	208	228	215	1	205	225
14	220	1	211	229	218	1	209	227
15	223	1	214	232	220	1	211	229
16	224	1	215	233	222	1	213	231
17	227	2	218	236	224	1	215	233
18	230	2	221	239	227	2	218	236
19	232	2	224	240	229	2	221	237
20	235	2	227	243	232	2	224	240
21	237	2	229	245	234	2	226	242
22	239	2	231	247	236	2	228	244
23	242	2	234	250	238	2	230	246
24	244	2	236	252	241	2	233	249
25	246	2	238	254	243	2	235	251
26	248	2	240	256	245	2	237	253
27	251	3	243	259	247	2	239	255
28	253	3	245	261	249	2	241	257
29	255	3	247	263	252	3	244	260
30	257	3	249	265	254	3	246	262
31	260	3	252	268	256	3	248	264
32	262	3	254	270	258	3	250	266
33	264	3	256	272	260	3	252	268
34	266	3	258	274	263	3	255	271
35	268	3	260	276	265	3	257	273
36	271	3	263	279	267	3	259	275
37	273	3	265	281	269	3	261	277
38	275	3	267	283	271	3	263	279
39	277	3	269	285	273	3	265	281
40	279	3	271	287	276	3	268	284
41	281	4	273	289	278	3	270	286
42	283	4	275	291	280	3	272	288
43	285	4	277	293	283	4	275	291
44	288	4	280	296	285	4	277	293
45	290	4	282	298	287	4	279	295
46	292	4	284	300	289	4	281	297
47	294	4	286	300	292	4	284	300
48	296	4	288	300	294	4	286	300
49	298	4	290	300	296	4	287	300
50	300	4	292	300	299	4	290	300
51	300	4	291	300	300	4	291	300
52	300	4	291	300	300	4	291	300
53	300	4	291	300	300	4	291	300
54	300	4	291	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300
62	300	4	290	300	300	4	290	300
63	300	4	290	300	300	4	290	300
64	300	4	290	300	300	4	290	300

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
65	300	4	290	300	300	4	290	300
66	300	4	290	300	300	4	290	300

Table M-8. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 3

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	203	1	200	213
9	202	1	200	212	205	1	200	215
10	205	1	200	215	208	1	200	218
11	207	1	200	217	210	1	200	220
12	209	1	200	219	213	1	203	223
13	212	1	202	222	215	1	205	225
14	214	1	204	224	218	1	209	227
15	217	1	208	226	220	1	211	229
16	219	1	210	228	223	1	214	232
17	221	1	212	230	225	2	216	234
18	224	1	215	233	228	2	219	237
19	226	2	217	235	230	2	221	239
20	229	2	220	238	233	2	224	242
21	231	2	222	240	235	2	227	243
22	234	2	225	243	238	2	230	246
23	236	2	227	245	240	2	232	248
24	239	2	231	247	242	2	234	250
25	241	2	233	249	245	2	237	253
26	244	2	236	252	247	2	239	255
27	247	2	239	255	249	2	241	257
28	249	2	241	257	252	3	244	260
29	252	3	244	260	255	3	247	263
30	254	3	246	262	257	3	249	265
31	257	3	249	265	259	3	251	267
32	259	3	251	267	262	3	254	270
33	262	3	254	270	264	3	256	272
34	264	3	256	272	266	3	258	274
35	267	3	259	275	269	3	261	277
36	269	3	260	278	271	3	263	279
37	272	3	263	281	274	3	266	282
38	274	3	265	283	276	3	267	285
39	277	3	268	286	278	3	269	287
40	279	3	270	288	281	3	272	290
41	282	3	273	291	283	3	274	292
42	284	3	275	293	285	3	276	294
43	286	3	277	295	288	4	279	297
44	289	4	280	298	290	4	281	299
45	291	4	281	300	292	4	282	300
46	293	4	283	300	295	4	285	300
47	296	4	286	300	297	4	287	300

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
48	298	4	288	300	299	4	289	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-9. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 4

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	202	1	200	212
11	202	1	200	212	205	1	200	215
12	204	1	200	214	207	1	200	217
13	207	1	200	217	210	1	200	220
14	209	1	200	219	212	1	202	222
15	211	1	201	221	215	1	205	225
16	214	1	204	224	218	1	208	228
17	216	1	206	226	220	1	210	230
18	219	1	209	229	223	1	214	232
19	222	1	212	232	225	2	216	234
20	224	1	215	233	228	2	219	237
21	227	2	218	236	230	2	221	239
22	229	2	220	238	233	2	224	242
23	232	2	223	241	235	2	226	244
24	234	2	225	243	238	2	229	247
25	237	2	228	246	240	2	231	249
26	239	2	230	248	243	2	234	252
27	242	2	233	251	245	2	236	254
28	245	2	236	254	248	2	239	257
29	247	2	238	256	251	3	242	260
30	249	2	240	258	253	3	244	262
31	253	3	244	262	256	3	247	265
32	255	3	246	264	258	3	249	267
33	258	3	249	267	261	3	252	270
34	261	3	252	270	263	3	254	272
35	263	3	254	272	266	3	257	275
36	266	3	257	275	268	3	259	277

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
37	269	3	260	278	271	3	262	280
38	272	3	262	282	274	3	265	283
39	274	3	264	284	276	3	266	286
40	277	3	267	287	279	3	269	289
41	280	3	270	290	281	3	271	291
42	283	3	273	293	284	3	274	294
43	285	3	275	295	287	3	277	297
44	288	3	278	298	289	4	279	299
45	291	4	281	300	292	4	282	300
46	294	4	284	300	295	4	285	300
47	296	4	286	300	297	4	287	300
48	299	4	289	300	300	4	290	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-10. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 5

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	201	1	200	211	202	1	200	212
13	204	1	200	214	205	1	200	215
14	207	1	200	217	208	1	200	218
15	210	1	200	220	211	1	201	221
16	213	1	203	223	214	1	204	224
17	216	1	206	226	216	1	206	226
18	219	1	209	229	219	1	209	229
19	222	1	212	232	222	1	212	232
20	224	1	214	234	225	2	215	235
21	227	2	217	237	228	2	218	238
22	230	2	220	240	231	2	221	241
23	233	2	223	243	233	2	223	243
24	236	2	226	246	236	2	226	246
25	239	2	229	249	239	2	229	249

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
26	241	2	231	251	242	2	232	252
27	244	2	234	254	244	2	234	254
28	247	2	237	257	247	2	237	257
29	249	2	239	259	249	2	239	259
30	253	3	243	263	253	3	243	263
31	255	3	245	265	255	3	245	265
32	258	3	248	268	258	3	248	268
33	261	3	251	271	261	3	251	271
34	264	3	254	274	263	3	253	273
35	266	3	256	276	266	3	256	276
36	269	3	259	279	269	3	259	279
37	272	3	262	282	271	3	261	281
38	274	3	264	284	274	3	264	284
39	277	3	267	287	277	3	267	287
40	280	3	270	290	279	3	269	289
41	282	3	272	292	282	3	272	292
42	285	3	275	295	284	3	274	294
43	288	4	278	298	287	4	277	297
44	290	4	280	300	290	4	280	300
45	293	4	283	300	292	4	282	300
46	295	4	285	300	295	4	285	300
47	298	4	288	300	297	4	287	300
48	300	4	290	300	300	4	290	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-11. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 6

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	201	1	200	211	200	1	200	210
13	205	1	200	215	203	1	200	213
14	208	1	200	218	206	1	200	216

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
15	212	1	202	222	209	1	200	219
16	215	1	205	225	212	1	202	222
17	218	1	208	228	215	1	205	225
18	222	1	212	232	219	1	209	229
19	224	1	214	234	222	1	212	232
20	228	2	218	238	224	1	214	234
21	231	2	221	241	228	2	218	238
22	234	2	224	244	231	2	221	241
23	237	2	227	247	234	2	224	244
24	240	2	230	250	236	2	226	246
25	243	2	233	253	239	2	229	249
26	246	2	236	256	242	2	232	252
27	249	2	239	259	245	2	235	255
28	252	3	242	262	248	2	238	258
29	255	3	245	265	251	3	241	261
30	258	3	248	268	254	3	244	264
31	261	3	251	271	256	3	246	266
32	264	3	254	274	259	3	249	269
33	266	3	256	276	262	3	252	272
34	269	3	259	279	265	3	255	275
35	272	3	262	282	267	3	257	277
36	275	3	265	285	270	3	260	280
37	277	3	267	287	273	3	263	283
38	280	3	270	290	276	3	266	286
39	283	3	273	293	278	3	268	288
40	286	3	276	296	281	3	271	291
41	288	3	278	298	284	3	274	294
42	291	4	281	300	286	3	276	296
43	293	4	283	300	289	4	279	299
44	296	4	286	300	292	4	282	300
45	299	4	289	300	295	4	285	300
46	300	4	290	300	297	4	287	300
47	300	4	290	300	300	4	290	300
48	300	4	290	300	300	4	290	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-12. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 7

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected Score Range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected Score Range	
			Low	High			Low	High
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	201	1	200	211	201	1	200	211
12	204	1	200	214	204	1	200	214
13	206	1	200	216	207	1	200	217
14	209	1	200	219	209	1	200	219
15	212	1	202	222	212	1	202	222
16	214	1	204	224	215	1	205	225
17	217	1	207	227	218	1	208	228
18	220	1	210	230	220	1	210	230
19	222	1	212	232	223	1	213	233
20	224	1	214	234	226	2	216	236
21	227	2	217	237	228	2	218	238
22	230	2	221	239	231	2	221	241
23	233	2	224	242	234	2	224	244
24	235	2	226	244	236	2	227	245
25	238	2	229	247	239	2	230	248
26	241	2	232	250	242	2	233	251
27	243	2	234	252	244	2	235	253
28	246	2	237	255	247	2	238	256
29	248	2	239	257	249	2	240	258
30	251	3	242	260	252	3	243	261
31	253	3	244	262	255	3	246	264
32	256	3	247	265	257	3	248	266
33	259	3	250	268	260	3	251	269
34	261	3	252	270	263	3	254	272
35	264	3	255	273	265	3	256	274
36	266	3	257	275	268	3	258	278
37	269	3	259	279	270	3	260	280
38	271	3	261	281	273	3	263	283
39	274	3	264	284	275	3	265	285
40	276	3	266	286	278	3	268	288
41	279	3	269	289	280	3	270	290
42	281	3	271	291	283	3	273	293
43	284	3	274	294	285	3	275	295
44	286	3	276	296	287	3	277	297
45	289	4	279	299	290	4	280	300
46	291	4	281	300	293	4	283	300
47	294	4	284	300	295	4	285	300
48	296	4	286	300	297	4	287	300
49	298	4	288	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-13. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 8

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	202	1	200	212	200	1	200	210
15	205	1	200	215	200	1	200	210
16	209	1	200	219	203	1	200	213
17	212	1	202	222	207	1	200	217
18	215	1	205	225	210	1	200	220
19	218	1	208	228	213	1	203	223
20	222	1	212	232	216	1	206	226
21	224	1	214	234	219	1	209	229
22	228	2	218	238	222	1	212	232
23	231	2	221	241	226	2	216	236
24	234	2	224	244	229	2	219	239
25	237	2	227	247	232	2	222	242
26	240	2	230	250	235	2	225	245
27	243	2	233	253	238	2	228	248
28	246	2	236	256	241	2	231	251
29	249	2	239	259	244	2	234	254
30	252	3	242	262	247	2	237	257
31	254	3	244	264	249	2	239	259
32	257	3	247	267	253	3	243	263
33	260	3	250	270	255	3	245	265
34	263	3	253	273	258	3	248	268
35	265	3	255	275	261	3	251	271
36	268	3	258	278	264	3	254	274
37	271	3	261	281	267	3	257	277
38	273	3	263	283	270	3	260	280
39	276	3	266	286	272	3	262	282
40	279	3	269	289	275	3	265	285
41	281	3	271	291	278	3	268	288
42	284	3	274	294	281	3	271	291
43	286	3	276	296	283	3	273	293
44	289	4	279	299	286	3	276	296
45	291	4	281	300	289	4	279	299
46	294	4	284	300	292	4	282	300
47	296	4	286	300	294	4	284	300
48	299	4	289	300	297	4	287	300
49	300	4	290	300	300	4	290	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-14. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 10

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	210	200	1	200	210
15	200	1	200	210	200	1	200	210
16	200	1	200	210	200	1	200	210
17	202	1	200	212	200	1	200	210
18	206	1	200	216	200	1	200	210
19	209	1	200	219	202	1	200	212
20	213	1	203	223	205	1	200	215
21	216	1	206	226	208	1	200	218
22	219	1	209	229	212	1	202	222
23	223	1	213	233	215	1	205	225
24	226	2	216	236	218	1	208	228
25	229	2	219	239	222	1	212	232
26	233	2	223	243	224	1	214	234
27	236	2	226	246	228	2	218	238
28	239	2	229	249	231	2	221	241
29	242	2	232	252	235	2	225	245
30	245	2	235	255	238	2	228	248
31	248	2	238	258	241	2	231	251
32	251	3	241	261	244	2	234	254
33	255	3	245	265	247	2	237	257
34	258	3	248	268	251	3	241	261
35	261	3	251	271	254	3	244	264
36	264	3	254	274	257	3	247	267
37	267	3	257	277	260	3	250	270
38	270	3	260	280	263	3	253	273
39	273	3	263	283	267	3	257	277
40	276	3	266	286	270	3	260	280
41	279	3	269	289	273	3	263	283
42	281	3	271	291	276	3	266	286
43	284	3	274	294	280	3	270	290

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
44	287	3	277	297	283	3	273	293
45	290	4	280	300	286	3	276	296
46	293	4	283	300	288	3	278	298
47	296	4	286	300	292	4	282	300
48	299	4	289	300	296	4	286	300
49	300	4	290	300	299	4	289	300
50	300	4	290	300	300	4	290	300
51	300	4	290	300	300	4	290	300
52	300	4	290	300	300	4	290	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300

Table M-15. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 4

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	209	200	1	200	209
12	200	1	200	209	200	1	200	209
13	200	1	200	209	202	1	200	211
14	200	1	200	209	205	1	200	213
15	203	1	200	211	208	1	200	216
16	205	1	200	213	210	1	202	218
17	208	1	200	216	213	1	205	221
18	210	1	202	218	215	1	207	223
19	213	1	205	221	217	1	209	225
20	215	1	207	223	220	1	212	228
21	217	1	209	225	222	1	215	229
22	219	1	211	227	224	1	217	231
23	222	1	214	230	226	2	219	233
24	224	1	217	231	228	2	221	235
25	226	2	219	233	230	2	223	237
26	228	2	221	235	232	2	225	239
27	230	2	223	237	234	2	227	241
28	232	2	225	239	236	2	229	243
29	234	2	227	241	238	2	231	245
30	236	2	229	243	240	2	233	247
31	238	2	231	245	242	2	235	249
32	240	2	233	247	243	2	236	250

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
33	242	2	235	249	245	2	238	252
34	244	2	237	251	247	2	240	254
35	246	2	239	253	249	2	242	256
36	248	2	241	255	251	3	244	258
37	249	2	242	256	253	3	246	260
38	252	3	245	259	255	3	248	262
39	254	3	247	261	257	3	250	264
40	256	3	248	264	259	3	252	266
41	258	3	250	266	261	3	253	269
42	261	3	253	269	264	3	256	272
43	263	3	255	271	266	3	258	274
44	265	3	257	273	268	3	260	276
45	268	3	260	276	271	3	263	279
46	270	3	262	278	273	3	265	281
47	273	3	265	281	276	3	267	285
48	276	3	267	285	279	3	270	288
49	278	3	269	287	282	4	273	291
50	281	3	272	290	285	4	276	294
51	285	4	276	294	288	4	278	298
52	288	4	278	298	292	4	282	300
53	292	4	282	300	296	4	286	300
54	297	4	287	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300

Table M-16. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 8

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	201	1	200	211
13	201	1	200	211	204	1	200	213
14	204	1	200	214	207	1	200	216
15	208	1	200	218	210	1	201	219
16	211	1	202	220	213	1	204	222
17	213	1	204	222	215	1	206	224
18	216	1	207	225	218	1	209	227
19	219	1	210	228	220	1	212	228
20	222	1	213	231	222	1	214	230

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
21	224	1	215	233	224	1	216	232
22	227	2	218	236	227	2	219	235
23	229	2	220	238	229	2	221	237
24	232	2	223	241	231	2	223	239
25	234	2	226	242	234	2	226	242
26	237	2	229	245	236	2	228	244
27	239	2	231	247	238	2	230	246
28	241	2	233	249	240	2	232	248
29	244	2	236	252	242	2	234	250
30	246	2	238	254	244	2	236	252
31	248	2	240	256	246	2	238	254
32	251	3	243	259	248	2	240	256
33	253	3	245	261	249	2	241	257
34	255	3	247	263	252	3	244	260
35	258	3	250	266	254	3	246	262
36	260	3	252	268	256	3	248	264
37	262	3	254	270	258	3	250	266
38	265	3	257	273	260	3	252	268
39	267	3	258	276	262	3	254	270
40	270	3	261	279	264	3	256	272
41	272	3	263	281	266	3	258	274
42	275	3	266	284	268	3	260	276
43	277	3	268	286	271	3	263	279
44	280	3	271	289	273	3	264	282
45	282	3	273	291	276	3	267	285
46	286	4	277	295	278	3	269	287
47	289	4	279	299	281	3	272	290
48	292	4	282	300	284	4	275	293
49	295	4	285	300	287	4	277	297
50	299	4	289	300	290	4	280	300
51	300	4	290	300	294	4	284	300
52	300	4	290	300	298	4	288	300
53	300	4	290	300	300	4	290	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300

Table M-17. 2010–11 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 10

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
0	200	1	200	210	200	1	200	210
1	200	1	200	210	200	1	200	210
2	200	1	200	210	200	1	200	210
3	200	1	200	210	200	1	200	210
4	200	1	200	210	200	1	200	210
5	200	1	200	210	200	1	200	210
6	200	1	200	210	200	1	200	210
7	200	1	200	210	200	1	200	210
8	200	1	200	210	200	1	200	210

continued

Raw score	2010–11				2009–10			
	Scaled score	Performance level	Expected score range		Scaled score	Performance level	Expected score range	
			Low	High			Low	High
9	200	1	200	210	200	1	200	210
10	200	1	200	210	200	1	200	210
11	200	1	200	210	200	1	200	210
12	200	1	200	210	200	1	200	210
13	200	1	200	210	200	1	200	210
14	200	1	200	209	200	1	200	209
15	201	1	200	210	201	1	200	210
16	204	1	200	213	203	1	200	212
17	207	1	200	216	206	1	200	215
18	209	1	200	218	209	1	200	218
19	212	1	203	221	211	1	203	219
20	214	1	206	222	214	1	206	222
21	217	1	209	225	216	1	208	224
22	219	1	211	227	218	1	210	226
23	221	1	213	229	221	1	213	229
24	224	1	216	232	223	1	215	231
25	226	2	218	234	224	1	216	232
26	228	2	220	236	227	2	219	235
27	230	2	222	238	229	2	221	237
28	233	2	225	241	231	2	223	239
29	235	2	227	243	234	2	226	242
30	237	2	229	245	236	2	228	244
31	239	2	231	247	238	2	230	246
32	241	2	233	249	240	2	232	248
33	244	2	236	252	242	2	234	250
34	246	2	238	254	244	2	236	252
35	248	2	240	256	246	2	238	254
36	250	3	242	258	248	2	240	256
37	253	3	245	261	251	3	243	259
38	255	3	247	263	253	3	245	261
39	257	3	249	265	255	3	247	263
40	260	3	252	268	257	3	249	265
41	262	3	254	270	260	3	252	268
42	265	3	256	274	262	3	254	270
43	267	3	258	276	265	3	256	274
44	270	4	261	279	267	3	258	276
45	273	4	264	282	270	4	261	279
46	275	4	266	284	273	4	264	282
47	278	4	269	287	276	4	267	285
48	282	4	272	292	279	4	269	289
49	285	4	275	295	282	4	272	292
50	288	4	278	298	286	4	276	296
51	292	4	282	300	290	4	280	300
52	296	4	286	300	294	4	284	300
53	300	4	290	300	298	4	288	300
54	300	4	290	300	300	4	290	300
55	300	4	290	300	300	4	290	300
56	300	4	290	300	300	4	290	300
57	300	4	290	300	300	4	290	300
58	300	4	290	300	300	4	290	300
59	300	4	290	300	300	4	290	300
60	300	4	290	300	300	4	290	300
61	300	4	290	300	300	4	290	300

Appendix N—CLASSICAL RELIABILITY

Table N-1. 2010–11 MontCAS: Subgroup Reliabilities – Mathematics

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	Special Education	1,222	66	35.28	13.15	0.92	3.69
	Title 1	4,469	66	41.39	12.11	0.91	3.60
	Low Income	5,168	66	40.70	11.68	0.90	3.62
	American Indian	1,439	66	36.25	11.72	0.90	3.64
	Asian	101	66	48.16	10.32	0.89	3.43
	Hispanic	371	66	41.47	11.40	0.90	3.58
	Black or African American	124	66	40.21	12.08	0.91	3.67
	White, Non-Hispanic	8,421	66	45.49	10.94	0.90	3.49
	Native Hawaiian/Other Pacific Islander	36	66	45.03	10.79	0.89	3.55
	Female	5,111	66	43.75	11.41	0.90	3.53
	Male	5,381	66	44.32	11.65	0.91	3.52
	Limited English Proficient	302	66	31.82	11.48	0.89	3.72
	Migrant	26	66	37.54	11.71	0.90	3.63
	Plan 504	28	66	43.57	12.44	0.92	3.58
	All Students	10,492	66	44.04	11.54	0.91	3.53
4	Special Education	1,226	66	32.92	12.88	0.91	3.77
	Title 1	4,490	66	39.82	12.17	0.91	3.67
	Low Income	5,010	66	39.55	12.01	0.91	3.68
	American Indian	1,329	66	34.86	12.09	0.90	3.75
	Asian	103	66	47.97	10.83	0.90	3.41
	Hispanic	404	66	40.33	11.30	0.90	3.65
	Black or African American	151	66	39.16	12.11	0.91	3.66
	White, Non-Hispanic	8,519	66	44.56	11.11	0.90	3.52
	Native Hawaiian/Other Pacific Islander	38	66	42.26	13.25	0.93	3.50
	Female	5,186	66	43.05	11.55	0.90	3.56
	Male	5,358	66	43.20	11.93	0.91	3.57
	Limited English Proficient	238	66	29.95	11.25	0.89	3.78
	Migrant	27	66	38.78	10.29	0.86	3.80
	Plan 504	56	66	43.68	11.25	0.90	3.53
	All Students	10,544	66	43.12	11.74	0.91	3.57
5	Special Education	1,212	66	29.26	12.22	0.90	3.79
	Title 1	4,351	66	37.89	12.84	0.92	3.74
	Low Income	4,769	66	37.26	12.56	0.91	3.75
	American Indian	1,283	66	32.89	12.64	0.91	3.78
	Asian	96	66	45.64	12.67	0.92	3.47
	Hispanic	382	66	36.98	12.07	0.90	3.73
	Black or African American	125	66	38.24	12.09	0.91	3.72

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
5	White,Non-Hispanic	8,442	66	42.52	11.89	0.91	3.64
	Native Hawaiian/Other Pacific Islander	38	66	42.26	12.79	0.92	3.63
	Female	4,993	66	40.71	12.13	0.91	3.68
	Male	5,373	66	41.46	12.73	0.92	3.65
	Limited English Proficient	257	66	25.85	10.64	0.88	3.71
	Migrant	30	66	34.40	12.93	0.91	3.82
	Plan 504	64	66	39.11	11.15	0.88	3.87
	All Students	10,366	66	41.10	12.45	0.91	3.67
6	Special Education	1,106	66	24.96	10.67	0.89	3.59
	Title 1	3,978	66	33.86	12.21	0.91	3.69
	Low Income	4,655	66	33.25	11.62	0.90	3.68
	American Indian	1,324	66	29.30	11.05	0.89	3.66
	Asian	94	66	41.39	12.37	0.91	3.62
	Hispanic	363	66	33.99	11.63	0.90	3.68
	Black or African American	145	66	34.94	12.07	0.91	3.66
	White,Non-Hispanic	8,568	66	38.65	11.74	0.90	3.66
	Native Hawaiian/Other Pacific Islander	28	66	41.04	10.58	0.89	3.49
	Female	5,072	66	37.07	11.59	0.90	3.65
	Male	5,450	66	37.50	12.52	0.91	3.68
	Limited English Proficient	217	66	23.97	9.17	0.84	3.61
	Migrant	22	66	34.82	9.97	0.86	3.75
	Plan 504	73	66	32.36	12.40	0.91	3.71
	All Students	10,522	66	37.29	12.08	0.91	3.67
7	Special Education	1,123	66	24.91	10.63	0.88	3.64
	Title 1	3,823	66	34.40	12.38	0.91	3.61
	Low Income	4,462	66	34.05	12.10	0.91	3.63
	American Indian	1,219	66	29.50	11.53	0.90	3.66
	Asian	92	66	42.07	13.22	0.93	3.47
	Hispanic	363	66	34.56	11.89	0.91	3.61
	Black or African American	118	66	32.37	10.30	0.87	3.69
	White,Non-Hispanic	8,662	66	39.82	12.03	0.91	3.52
	Native Hawaiian/Other Pacific Islander	28	66	41.93	9.56	0.87	3.42
	Female	5,093	66	38.55	12.03	0.91	3.53
	Male	5,389	66	38.21	12.81	0.92	3.55
	Limited English Proficient	221	66	21.48	8.78	0.83	3.62
	Migrant	17	66	33.29	9.45	0.84	3.78
	Plan 504	91	66	37.08	10.53	0.88	3.58
	All Students	10,482	66	38.38	12.44	0.92	3.55

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
8	Special Education	1,086	66	23.51	9.73	0.86	3.68
	Title 1	3,847	66	33.59	12.16	0.90	3.79
	Low Income	4,282	66	32.75	11.73	0.90	3.80
	American Indian	1,203	66	28.89	10.97	0.88	3.77
	Asian	105	66	41.96	11.62	0.90	3.68
	Hispanic	368	66	33.45	11.68	0.89	3.81
	Black or African American	134	66	31.40	11.38	0.89	3.78
	White,Non-Hispanic	8,607	66	38.54	12.02	0.90	3.73
	Native Hawaiian/Other Pacific Islander	31	66	35.48	12.02	0.90	3.85
	Female	5,078	66	37.43	12.02	0.90	3.76
	Male	5,370	66	36.96	12.59	0.91	3.73
	Limited English Proficient	226	66	21.94	8.69	0.82	3.65
	Migrant	13	66	32.31	8.68	0.82	3.71
	Plan 504	106	66	32.67	11.36	0.89	3.79
	All Students	10,448	66	37.19	12.32	0.91	3.75
10	Special Education	924	66	19.12	7.59	0.79	3.47
	Title 1	2,914	66	27.40	12.01	0.91	3.68
	Low Income	3,371	66	26.51	11.06	0.89	3.67
	American Indian	1,030	66	23.27	9.78	0.87	3.58
	Asian	111	66	34.93	13.55	0.92	3.75
	Hispanic	304	66	28.41	10.58	0.88	3.71
	Black or African American	89	66	26.53	10.50	0.88	3.67
	White,Non-Hispanic	8,494	66	32.46	12.30	0.91	3.74
	Native Hawaiian/Other Pacific Islander	31	66	28.32	13.75	0.93	3.74
	Female	4,872	66	30.65	11.92	0.90	3.75
	Male	5,187	66	32.03	12.73	0.92	3.71
	Limited English Proficient	131	66	17.59	7.32	0.79	3.37
	Migrant	15	66	31.07	8.98	0.83	3.69
	Plan 504	138	66	29.17	11.64	0.90	3.71
	All Students	10,059	66	31.36	12.36	0.91	3.73

Table N-2. 2010–11 MontCAS: Subgroup Reliabilities – Reading

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	Special Education	1,192	60	32.01	11.96	0.92	3.32
	Title 1	4,740	60	37.76	10.74	0.91	3.20
	Low Income	5,145	60	37.51	10.59	0.91	3.21
	American Indian	1,439	60	33.48	10.72	0.90	3.32
	Asian	101	60	43.72	9.23	0.89	3.03
	Hispanic	369	60	38.11	10.70	0.91	3.21
	Black or African American	123	60	37.63	10.85	0.91	3.25
	White, Non-Hispanic	8,394	60	41.89	9.74	0.90	3.06
	Native Hawaiian/Other Pacific Islander	36	60	40.56	8.50	0.87	3.09
	Female	5,103	60	41.52	10.10	0.91	3.10
	Male	5,359	60	39.65	10.50	0.91	3.12
	Limited English Proficient	300	60	28.50	10.04	0.88	3.43
	Migrant	26	60	35.81	8.78	0.86	3.23
	Plan 504	28	60	37.79	12.20	0.93	3.11
	All Students	10,462	60	40.56	10.35	0.91	3.11
4	Special Education	1,205	60	31.84	10.96	0.91	3.33
	Title 1	4,558	60	38.04	10.50	0.91	3.17
	Low Income	4,993	60	37.60	10.40	0.91	3.18
	American Indian	1,320	60	33.49	10.39	0.90	3.31
	Asian	103	60	43.14	10.50	0.92	2.93
	Hispanic	403	60	39.00	9.86	0.90	3.13
	Black or African American	151	60	38.54	10.85	0.92	3.09
	White, Non-Hispanic	8,510	60	42.15	9.36	0.90	2.98
	Native Hawaiian/Other Pacific Islander	37	60	37.65	11.60	0.93	3.11
	Female	5,184	60	41.79	9.77	0.91	3.00
	Male	5,340	60	40.01	10.12	0.91	3.06
	Limited English Proficient	235	60	28.18	9.53	0.87	3.42
	Migrant	27	60	38.00	9.77	0.89	3.18
	Plan 504	56	60	39.93	9.57	0.90	3.08
	All Students	10,524	60	40.89	9.99	0.91	3.03
5	Special Education	1,197	60	31.62	10.95	0.90	3.41
	Title 1	4,309	60	39.34	10.61	0.91	3.22
	Low Income	4,764	60	38.86	10.47	0.90	3.24
	American Indian	1,277	60	35.04	11.06	0.91	3.34
	Asian	96	60	43.96	10.36	0.92	2.97
	Hispanic	381	60	38.64	10.47	0.90	3.24
	Black or African American	123	60	40.91	10.17	0.90	3.18

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
5	White,Non-Hispanic	8,441	60	43.24	9.22	0.89	3.06
	Native Hawaiian/Other Pacific Islander	38	60	42.82	9.71	0.90	3.10
	Female	4,992	60	43.10	9.56	0.90	3.08
	Male	5,364	60	41.04	10.16	0.91	3.11
	Limited English Proficient	253	60	27.22	9.36	0.86	3.54
	Migrant	30	60	38.47	11.92	0.92	3.32
	Plan 504	63	60	41.57	9.45	0.89	3.15
	All Students	10,356	60	42.04	9.93	0.90	3.11
6	Special Education	1,104	60	28.83	10.35	0.89	3.41
	Title 1	3,984	60	37.40	10.66	0.91	3.23
	Low Income	4,657	60	36.96	10.35	0.90	3.24
	American Indian	1,324	60	33.17	10.50	0.90	3.35
	Asian	94	60	43.01	9.50	0.90	3.04
	Hispanic	363	60	37.91	9.94	0.90	3.20
	Black or African American	145	60	39.27	9.98	0.90	3.17
	White,Non-Hispanic	8,569	60	41.73	9.47	0.89	3.07
	Native Hawaiian/Other Pacific Islander	28	60	42.43	9.07	0.89	3.04
	Female	5,073	60	41.81	9.64	0.90	3.08
	Male	5,450	60	39.28	10.27	0.91	3.12
	Limited English Proficient	217	60	26.22	9.03	0.85	3.48
	Migrant	22	60	35.59	8.07	0.82	3.43
	Plan 504	73	60	37.26	10.66	0.91	3.21
	All Students	10,523	60	40.50	10.05	0.90	3.12
7	Special Education	1,133	60	29.46	11.38	0.91	3.46
	Title 1	3,781	60	39.50	12.06	0.93	3.23
	Low Income	4,468	60	39.29	11.77	0.92	3.25
	American Indian	1,222	60	34.36	12.26	0.92	3.37
	Asian	92	60	44.62	11.36	0.93	2.97
	Hispanic	365	60	40.48	11.20	0.92	3.23
	Black or African American	118	60	40.08	10.85	0.91	3.18
	White,Non-Hispanic	8,670	60	44.57	10.28	0.91	3.01
	Native Hawaiian/Other Pacific Islander	28	60	46.32	8.34	0.87	2.96
	Female	5,101	60	44.75	10.44	0.92	3.01
	Male	5,394	60	41.73	11.47	0.93	3.10
	Limited English Proficient	220	60	22.66	8.98	0.85	3.51
	Migrant	17	60	37.29	9.23	0.87	3.35
	Plan 504	91	60	43.16	9.34	0.89	3.13
	All Students	10,495	60	43.20	11.08	0.92	3.08

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
8	Special Education	1,118	60	27.95	11.24	0.90	3.53
	Title 1	3,832	60	38.91	12.29	0.93	3.35
	Low Income	4,307	60	38.38	12.08	0.92	3.37
	American Indian	1,204	60	34.45	12.29	0.92	3.46
	Asian	106	60	44.09	11.24	0.92	3.17
	Hispanic	369	60	39.49	11.66	0.92	3.36
	Black or African American	136	60	37.80	12.38	0.93	3.39
	White,Non-Hispanic	8,638	60	43.81	10.78	0.92	3.13
	Native Hawaiian/Other Pacific Islander	31	60	40.48	11.82	0.92	3.33
	Female	5,091	60	44.47	10.76	0.92	3.10
	Male	5,393	60	40.64	11.76	0.92	3.22
	Limited English Proficient	226	60	24.27	10.08	0.87	3.57
	Migrant	13	60	35.54	8.69	0.83	3.57
	Plan 504	107	60	40.31	11.30	0.91	3.32
	All Students	10,484	60	42.50	11.45	0.92	3.19
10	Special Education	940	60	28.24	10.44	0.89	3.52
	Title 1	2,848	60	38.05	11.35	0.91	3.32
	Low Income	3,379	60	37.76	11.40	0.91	3.34
	American Indian	1,032	60	34.47	11.24	0.91	3.45
	Asian	111	60	42.80	10.94	0.92	3.14
	Hispanic	306	60	40.08	11.19	0.92	3.23
	Black or African American	89	60	39.88	10.56	0.90	3.29
	White,Non-Hispanic	8,508	60	42.86	10.22	0.91	3.14
	Native Hawaiian/Other Pacific Islander	31	60	38.84	12.32	0.92	3.38
	Female	4,878	60	43.55	10.00	0.90	3.12
	Male	5,199	60	40.30	11.07	0.92	3.20
	Limited English Proficient	131	60	22.89	8.91	0.84	3.55
	Migrant	15	60	41.73	10.57	0.91	3.18
	Plan 504	138	60	40.34	10.83	0.91	3.30
	All Students	10,077	60	41.87	10.69	0.91	3.18

Table N-3. 2010–11 MontCAS: Subgroup Reliabilities – Science

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
4	Special Education	1,253	61	33.44	10.25	0.88	3.57
	Title 1	13	61	32.69	9.67	0.87	3.50
	Low Income	5,030	61	36.50	9.74	0.87	3.54
	American Indian	1,331	61	31.91	9.53	0.86	3.60
	Asian	103	61	42.08	9.30	0.87	3.36
	Hispanic	407	61	36.82	9.12	0.85	3.56
	Black or African American	152	61	36.06	9.51	0.86	3.54
	White,Non-Hispanic	8,543	61	41.22	9.01	0.86	3.39
	Native Hawaiian/Other Pacific Islander	38	61	35.84	10.68	0.89	3.53
	Female	5,202	61	39.55	9.52	0.87	3.44
	Male	5,372	61	40.02	9.74	0.88	3.43
	Limited English Proficient	238	61	27.40	8.57	0.82	3.67
	Migrant	27	61	37.96	9.83	0.87	3.54
	Plan 504	57	61	40.81	9.56	0.88	3.35
	All Students	10,574	61	39.79	9.63	0.87	3.44
8	Special Education	1,131	61	25.03	9.09	0.85	3.52
	Title 1	28	61	29.79	8.89	0.84	3.56
	Low Income	4,307	61	31.18	9.95	0.87	3.52
	American Indian	1,201	61	27.12	9.37	0.86	3.51
	Asian	106	61	37.83	9.65	0.87	3.47
	Hispanic	369	61	31.98	9.67	0.87	3.54
	Black or African American	136	61	30.60	10.30	0.88	3.53
	White,Non-Hispanic	8,640	61	36.13	9.52	0.87	3.46
	Native Hawaiian/Other Pacific Islander	31	61	33.77	10.69	0.89	3.46
	Female	5,083	61	34.86	9.50	0.87	3.45
	Male	5,400	61	34.91	10.40	0.89	3.49
	Limited English Proficient	225	61	20.82	7.15	0.76	3.49
	Migrant	13	61	28.62	8.17	0.79	3.72
	Plan 504	105	61	34.62	9.89	0.87	3.55
	All Students	10,483	61	34.89	9.97	0.88	3.48
10	Special Education	954	61	23.65	8.71	0.84	3.51
	Title 1	63	61	27.38	7.67	0.79	3.52
	Low Income	3,381	61	30.05	9.94	0.87	3.53
	American Indian	1,027	61	26.16	9.07	0.85	3.54
	Asian	111	61	37.15	10.77	0.90	3.45
	Hispanic	307	61	31.05	9.77	0.87	3.56
	Black or African American	89	61	31.16	9.87	0.87	3.52

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
10	White, Non-Hispanic	8,512	61	35.19	10.09	0.88	3.48
	Native Hawaiian/Other Pacific Islander	31	61	31.97	10.29	0.88	3.54
	Female	4,876	61	33.76	9.87	0.87	3.54
	Male	5,201	61	34.45	10.80	0.90	3.44
	Limited English Proficient	130	61	19.91	6.82	0.74	3.45
	Migrant	15	61	34.60	9.58	0.87	3.44
	Plan 504	138	61	33.26	10.74	0.89	3.51
	All Students	10,077	61	34.12	10.37	0.89	3.49

Table N-4. 2010–11 MontCAS: Reliabilities by Reporting Category -- Mathematics

Grade	Item reporting category*	Number of items	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	2	19	22	15.38	4.30	0.77	2.06
	3	8	8	4.40	2.03	0.67	1.17
	4	10	10	6.58	2.04	0.58	1.32
	5	10	10	7.46	1.87	0.61	1.17
	6	5	8	4.65	2.06	0.43	1.55
	7	8	8	5.57	1.74	0.62	1.07
4	2	19	22	14.67	4.39	0.77	2.09
	3	8	8	5.73	1.79	0.64	1.08
	4	7	10	5.89	2.27	0.49	1.62
	5	10	10	6.68	2.05	0.57	1.34
	6	8	8	5.10	1.66	0.52	1.15
	7	8	8	5.06	2.04	0.67	1.18
5	2	21	21	13.24	4.79	0.85	1.88
	3	8	8	5.25	1.86	0.60	1.18
	4	8	11	6.94	2.55	0.51	1.78
	5	8	8	4.48	2.01	0.60	1.26
	6	7	10	5.80	2.13	0.49	1.52
	7	8	8	5.39	1.71	0.52	1.18
6	2	18	21	11.08	4.50	0.79	2.08
	3	8	8	4.90	1.90	0.57	1.25
	4	8	11	4.85	2.39	0.55	1.61
	5	8	8	5.08	1.80	0.55	1.20
	6	10	10	6.04	2.23	0.60	1.41
	7	8	8	5.35	1.84	0.57	1.20
7	2	18	18	10.38	4.09	0.81	1.80
	3	5	8	4.61	1.91	0.49	1.36
	4	9	12	6.12	2.69	0.66	1.56
	5	8	8	4.09	1.85	0.55	1.24
	6	12	12	8.01	2.66	0.70	1.45
	7	8	8	5.16	1.72	0.57	1.13
8	2	18	18	9.63	3.81	0.76	1.85
	3	5	8	4.39	2.10	0.50	1.48
	4	9	12	6.35	2.77	0.60	1.75
	5	8	8	4.48	1.82	0.53	1.25
	6	12	12	7.64	2.59	0.67	1.48
	7	8	8	4.70	1.82	0.53	1.24
10	2	10	13	5.54	3.17	0.68	1.80
	3	11	11	5.25	2.35	0.63	1.42
	4	10	13	5.93	2.68	0.58	1.74
	5	8	8	4.06	2.09	0.64	1.26
	6	13	13	6.53	2.78	0.67	1.59
	7	8	8	4.05	1.75	0.51	1.23

*Please note: 2: Numbers and Operations; 3: Algebra; 4: Geometry; 5: Measurement; 6: Data Analysis, Statistics, and Probability; 7: Patterns, Relations, and Functions

Table N-5. 2010–11 MontCAS: Reliabilities by Reporting Category – Reading

Grade	Item reporting category	Number of items	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	1	18	24	15.64	4.36	0.80	1.94
	2	19	19	14.26	3.45	0.77	1.66
	4	9	9	5.90	1.96	0.59	1.25
	5	8	8	4.76	1.92	0.58	1.25
4	1	18	21	14.86	3.65	0.78	1.69
	2	16	19	12.29	3.22	0.72	1.71
	4	10	10	7.20	2.01	0.58	1.29
	5	10	10	6.55	2.46	0.72	1.30
5	1	20	23	16.55	3.87	0.76	1.91
	2	15	18	11.94	3.31	0.73	1.71
	4	10	10	7.25	2.08	0.64	1.24
	5	9	9	6.30	2.00	0.62	1.24
6	1	20	23	15.69	4.07	0.78	1.90
	2	17	17	12.32	3.26	0.75	1.62
	4	8	11	6.53	2.05	0.55	1.38
	5	9	9	5.96	2.02	0.61	1.26
7	1	16	19	13.28	3.90	0.79	1.80
	2	17	17	12.84	3.24	0.77	1.55
	4	11	11	8.57	2.34	0.73	1.22
	5	10	13	8.50	2.75	0.69	1.52
8	1	17	17	12.58	3.63	0.81	1.59
	2	14	17	11.84	3.47	0.74	1.76
	4	10	10	7.75	2.06	0.67	1.18
	5	13	16	10.33	3.52	0.75	1.77
10	1	17	17	11.65	3.39	0.74	1.73
	2	15	18	12.50	3.52	0.74	1.78
	4	15	15	11.42	2.92	0.75	1.47
	5	7	10	6.31	2.16	0.61	1.34

*Please note:

1. Students construct meaning as they comprehend, interpret, and respond to what they read.
2. Students apply a range of skills and strategies to read.
4. Students select, read, and respond to print and nonprint material for a variety of purposes.
5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences.

Table N-6. 2010–11 MontCAS: Reliabilities by Reporting Category – Science

<i>Grade</i>	<i>Item reporting category</i>	<i>Number of items</i>	<i>Raw score</i>			<i>Alpha</i>	<i>SEM</i>
			<i>Maximum</i>	<i>Mean</i>	<i>Standard deviation</i>		
4	1	11	14	8.50	3.11	0.64	1.86
	2	11	14	8.32	2.40	0.55	1.61
	3	14	14	9.78	2.61	0.64	1.56
	4	14	14	9.51	2.50	0.60	1.59
	5	3	3	1.94	0.89	0.26	0.76
	6	2	2	1.75	0.50	0.25	0.43
8	1	14	14	8.61	2.99	0.70	1.63
	2	11	14	7.14	2.45	0.58	1.59
	3	14	14	8.28	2.76	0.64	1.65
	4	11	14	7.37	2.70	0.54	1.83
	5	3	3	1.89	0.92	0.30	0.77
	6	2	2	1.60	0.58	0.18	0.52
10	1	14	14	8.64	2.89	0.68	1.63
	2	14	14	8.15	2.75	0.63	1.68
	3	11	14	6.79	2.65	0.61	1.65
	4	14	14	8.61	2.91	0.69	1.62
	5	1	4	1.54	1.15		
	6	1	1	0.38	0.48		

*Please note: 1: Science Investigations; 2: Physical Science; 3: Life Science; 4: Earth/Space Science; 5: Impact on Society; 6: Historical Development

Appendix O—INTERRATER AGREEMENT

Table O-1. 2010–11 MontCAS: Item Level Interrater Consistency Statistics by Subject and Grade

<i>Subject</i>	<i>Grade</i>	<i>Item</i>	<i>Number of score categories</i>	<i>Number of responses scored twice</i>	<i>Percent exact</i>	<i>Percent adjacent</i>	<i>Correlation</i>	<i>Percent of third scores</i>
Mathematics	3	139045	2	229	97.82	2.18	0.93	0.00
		139051	2	216	99.07	0.93	0.98	0.00
		138799	5	263	82.89	14.83	0.93	2.28
		77041	2	209	99.04	0.96	0.98	0.00
		43277	5	231	78.35	19.48	0.92	2.16
	4	140171	2	212	96.23	3.77	0.91	0.00
		76773	2	211	100.00	0.00	1.00	0.00
		140183	5	245	85.31	13.47	0.94	1.22
		76755	2	219	99.09	0.91	0.97	0.00
		205515	5	218	82.57	11.01	0.87	6.42
	5	34605	2	209	98.56	1.44	0.97	0.00
		77299	2	403	99.01	0.99	0.97	0.00
		206043	5	237	70.89	27.00	0.86	2.11
		62035	2	209	96.17	3.83	0.91	0.00
		206104	5	213	88.26	9.86	0.95	1.88
	6	141512	2	212	98.58	1.42	0.96	0.00
		77646	2	215	98.14	1.86	0.95	0.00
		43913	2	216	100.00	0.00	1.00	0.00
		63054	5	211	91.00	8.06	0.95	0.95
		206762	5	239	82.85	15.48	0.92	1.67
	7	43911	2	208	98.08	1.92	0.96	0.00
		61374	2	214	99.53	0.47	0.99	0.00
		86339	2	213	98.59	1.41	0.97	0.00
		142673	5	210	85.71	13.81	0.94	0.48
		43828	5	217	80.65	17.97	0.90	1.38
	8	243715	2	218	99.54	0.46	0.99	0.00
		243699	2	201	95.52	4.48	0.91	0.00
		243710	2	212	99.06	0.94	0.98	0.00
		87874	5	237	82.28	15.61	0.93	1.69
		205490	5	221	87.78	11.76	0.96	0.45
	10	77644	2	190	98.42	1.58	0.96	0.00
		145015	2	197	98.98	1.02	0.98	0.00
		77629	2	197	99.49	0.51	0.99	0.00
		62391	5	189	68.25	28.04	0.82	3.70
		59407	5	213	89.67	8.92	0.96	1.41
Reading	3	151240	5	229	61.57	37.55	0.76	0.87
		150991	5	216	66.20	33.33	0.84	0.46
	4	151772	5	218	69.72	27.52	0.68	2.75
		151659	5	213	63.85	34.27	0.70	1.88
	5	150456	5	208	58.17	38.94	0.72	2.88
		150518	5	210	65.24	32.86	0.75	1.90
	6	95469	5	216	71.76	28.24	0.81	0.00
		68294	5	211	63.51	35.07	0.73	1.42
	7	148887	5	211	57.35	41.71	0.76	0.95
		149016	5	237	63.71	35.02	0.83	0.84
	8	149368	5	248	56.85	36.69	0.75	6.45
		149193	5	222	61.71	36.94	0.81	1.35
	10	149566	5	189	58.73	40.21	0.77	1.06
		149648	5	202	58.91	40.10	0.82	0.99

continued

<i>Subject</i>	<i>Grade</i>	<i>Item</i>	<i>Number of score categories</i>	<i>Number of responses scored twice</i>	<i>Percent exact</i>	<i>Percent adjacent</i>	<i>Correlation</i>	<i>Percent of third scores</i>
Science	4	75757	5	212	69.81	25.94	0.87	4.25
		120312	5	215	74.42	21.40	0.80	4.19
	8	89534	5	208	77.88	21.63	0.89	0.48
		89958	5	210	88.57	10.95	0.96	0.48
	10	119960	5	198	54.04	38.38	0.72	7.58
		52953	5	198	72.22	25.76	0.75	2.02

Appendix P—DECISION ACCURACY AND CONSISTENCY RESULTS

**Table P-1. 2010–11 MontCAS: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—
Overall and Conditional on Performance Level**

<i>Subject</i>	<i>Grade</i>	<i>Overall</i>	<i>Kappa</i>	<i>Conditional on level</i>			
				<i>Novice</i>	<i>Nearing Proficiency</i>	<i>Proficient</i>	<i>Advanced</i>
Mathematics	3	0.78 (0.71)	0.58	0.82 (0.74)	0.61 (0.49)	0.77 (0.70)	0.89 (0.80)
	4	0.79 (0.71)	0.59	0.82 (0.74)	0.64 (0.53)	0.76 (0.69)	0.89 (0.81)
	5	0.79 (0.72)	0.59	0.80 (0.71)	0.63 (0.52)	0.77 (0.69)	0.90 (0.83)
	6	0.78 (0.70)	0.58	0.79 (0.69)	0.66 (0.55)	0.74 (0.66)	0.90 (0.83)
	7	0.79 (0.72)	0.60	0.81 (0.73)	0.62 (0.51)	0.74 (0.66)	0.91 (0.85)
	8	0.79 (0.71)	0.59	0.81 (0.71)	0.68 (0.58)	0.78 (0.71)	0.89 (0.80)
	10	0.77 (0.69)	0.56	0.68 (0.55)	0.70 (0.61)	0.77 (0.69)	0.90 (0.81)
Reading	3	0.84 (0.78)	0.65	0.78 (0.66)	0.71 (0.60)	0.80 (0.74)	0.91 (0.86)
	4	0.84 (0.77)	0.64	0.78 (0.66)	0.72 (0.62)	0.80 (0.74)	0.91 (0.85)
	5	0.85 (0.79)	0.63	0.79 (0.67)	0.67 (0.55)	0.77 (0.69)	0.93 (0.89)
	6	0.85 (0.79)	0.63	0.77 (0.65)	0.64 (0.52)	0.78 (0.72)	0.92 (0.88)
	7	0.86 (0.81)	0.67	0.79 (0.69)	0.66 (0.54)	0.79 (0.72)	0.94 (0.90)
	8	0.85 (0.79)	0.65	0.80 (0.70)	0.61 (0.50)	0.77 (0.70)	0.94 (0.90)
	10	0.83 (0.77)	0.62	0.81 (0.72)	0.62 (0.50)	0.76 (0.69)	0.92 (0.87)
Science	4	0.78 (0.69)	0.54	0.75 (0.61)	0.75 (0.68)	0.79 (0.73)	0.82 (0.65)
	8	0.78 (0.69)	0.54	0.78 (0.67)	0.70 (0.61)	0.80 (0.75)	0.82 (0.68)
	10	0.75 (0.66)	0.54	0.82 (0.74)	0.73 (0.65)	0.65 (0.55)	0.86 (0.74)

Table P-2. 2010–11 MontCAS: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Conditional on Cutpoint

Subject	Grade	Novice/Nearing Proficiency			Nearing Proficiency/Proficient			Proficient/Advanced		
		Accuracy (consistency)	False positive	False negative	Accuracy (consistency)	False positive	False negative	Accuracy (consistency)	False positive	False negative
Mathematics	3	0.95 (0.93)	0.02	0.03	0.92 (0.89)	0.04	0.04	0.91 (0.88)	0.06	0.03
	4	0.95 (0.94)	0.02	0.02	0.93 (0.90)	0.04	0.04	0.91 (0.87)	0.06	0.03
	5	0.96 (0.94)	0.02	0.02	0.93 (0.90)	0.04	0.03	0.91 (0.88)	0.06	0.03
	6	0.95 (0.93)	0.02	0.03	0.92 (0.88)	0.05	0.04	0.92 (0.88)	0.05	0.03
	7	0.95 (0.93)	0.02	0.03	0.92 (0.89)	0.04	0.04	0.92 (0.89)	0.05	0.03
	8	0.95 (0.94)	0.02	0.02	0.92 (0.88)	0.04	0.04	0.92 (0.89)	0.05	0.03
	10	0.93 (0.91)	0.03	0.04	0.91 (0.87)	0.05	0.04	0.93 (0.91)	0.04	0.02
Reading	3	0.98 (0.98)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.90 (0.87)	0.06	0.04
	4	0.98 (0.97)	0.01	0.01	0.95 (0.93)	0.02	0.02	0.90 (0.87)	0.05	0.04
	5	0.98 (0.98)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04
	6	0.98 (0.98)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04
	7	0.98 (0.97)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.92 (0.89)	0.05	0.03
	8	0.98 (0.96)	0.01	0.01	0.95 (0.94)	0.02	0.02	0.92 (0.88)	0.05	0.04
	10	0.97 (0.96)	0.01	0.01	0.95 (0.93)	0.03	0.03	0.91 (0.87)	0.05	0.04
Science	4	0.96 (0.95)	0.02	0.02	0.90 (0.86)	0.06	0.05	0.92 (0.89)	0.06	0.02
	8	0.95 (0.93)	0.02	0.03	0.90 (0.87)	0.05	0.04	0.92 (0.89)	0.05	0.02
	10	0.93 (0.90)	0.03	0.04	0.90 (0.86)	0.06	0.04	0.92 (0.89)	0.05	0.03

Appendix Q—REPORT SHELLS



MontCAS Criterion-Referenced Test (CRT) Student Report 2011

Letter from Superintendent

Dear Parents/Guardians:

The Montana Comprehensive Assessment System (MontCAS) Criterion-Referenced Test (CRT) is the state's measure of student performance on the state content standards which establish goals for what all students should know and be able to do.

The CRT assesses Reading and Math at grades 3-8 and 10. Students in grades 4, 8, and 10 are also assessed in Science. The assessment contains multiple-choice questions, math short answer questions, and constructed response items. The constructed response items give students the opportunity to explain answers and solve problems using multiple strategies.

This report shows how your student performed on the March 2011 CRT. The results of this standards-based assessment are reported in four performance levels: Advanced, Proficient, Nearing Proficiency, and Novice. While some students may not yet meet the standards, keep in mind that the standards are rigorous and challenging. Our long term goal is for all students to achieve these high standards so that Montana youth will be among the best educated in the world. The staff at your school will be able to provide further information about your student's performance on the CRT.

The CRT is only one measure of student performance and should be viewed in the context of the student's local programs and other measures. The CRT is required by the No Child Left Behind Act and is part of an ongoing statewide educational improvement process. I encourage you to contact your student's school to begin a conversation that will support your student's success.

Sincerely,

Denise Juneau
Montana Superintendent of Public Instruction
Montana Office of Public Instruction
PO Box 202501
Helena, Montana 59620-2501
<http://www.opi.mt.gov>

What can you do to help your student?

It is important to support your student in his or her studies now and throughout his or her future education.

Here are some tips for supporting your student in the completion of his or her schoolwork:

- Have regular discussions with your student's teacher(s) to see what you can do at home to support your student's work in school, such as making sure homework is done.
- Discuss with your student the subjects in which he or she needs improvement. Talk about whether there has been a noticeable improvement. If not, find out why.
- Ask your student to explain what he or she is studying. These conversations help you to follow your student's progress and help your student to remember what he or she has learned.
- Make sure your student gets enough rest, eats properly, and arrives at school on time every day. Send your student to school prepared to learn.

What is the MontCAS Criterion-Referenced Test (CRT)?

The Montana Comprehensive Assessment System (MontCAS) was developed in accordance with the following federal laws: Title 1 of the Elementary and Secondary Education Act (ESEA) of 1994, P. L. 103-382, and the No Child Left Behind Act (NCLB) of 2001.

The CRT test questions are based on, and aligned to, Montana's content standards, benchmarks, and grade-level expectations in Mathematics, Reading, and Science. Montana educators worked with the Montana Office of Public Instruction and Measured Progress to develop test questions that assess how well students have met Montana grade-level expectations for each

MontCAS CRT scores are intended to be useful indicators of the extent to which students have mastered the materials outlined in the Montana Mathematics, Reading, and Science content standards, benchmarks, and grade-level expectations.

Who must take the CRT?

All classroom students in grades 3-8 and 10 enrolled for 180 hours or more in an accredited public or private Montana school are required to participate.

What subjects were tested in spring 2011?

Mathematics	Grades 3-8 and 10
Reading	Grades 3-8 and 10
Science	Grades 4, 8, and 10

What types of test questions are on the CRT?

- Multiple-choice questions: Students choose the correct answer from four options and receive one point for each correct answer and zero points for an incorrect answer.
- Constructed-response questions: Depending on the subject tested, students are asked to explain and/or make a chart, table, diagram, illustration, or graph to support their answer. Each answer receives zero to four points.
- Short-answer questions (Mathematics tests only): Students give a brief response, which is usually a number or short statement. Students receive one point for a correct answer and zero points for an incorrect answer.

How are the CRT results used?

MontCAS CRT test results are used for the following purposes:

- to assist educators in planning improvements to curriculum and instruction
- to determine whether schools are helping their students meet the state content standards

Where can you find more information?

Where can you find more information:
<https://data.opi.mt.gov/opireportingcenter>

Montana requirements for the participation of students with disabilities on the CRT:
http://www.opi.mt.gov/Curriculum/MontCAS/#p7GPd_7

OPI contact: Judy Snow, State Assessment
Director, 406-444-3656,
jsnow@mt.gov

Your student's performance level and score in each content area

Display of scores and probable range of scores

In the figure below your student's performance is displayed. For each subject, the left column lists the possible performance levels with the scores needed to achieve those levels. The center column is your student's performance where the black bar is their score and the small grey bar is the range of scores they might have achieved had they taken the test multiple times. The right hand column is the percentage of students that achieved each performance level on the CRT across the state.

Example: Your child's score → 240
← Range of likely scores if your child took the test many times

Mathematics			Reading			Science		
Performance Levels	Student	State percentage	Performance Levels	Student	State percentage	Performance Levels	Student	State percentage
Advanced 291-300	300	31%	Advanced 289-300	300	44%	Advanced 282-300		13%
Proficient 250-290		39%	Proficient 250-288		39%	Proficient 250-281	268	49%
Nearing Proficiency 225-249		17%	Nearing Proficiency 225-249		12%	Nearing Proficiency 225-249		30%
Novice 200-224		13%	Novice 200-224		4%	Novice 200-224		8%

Your student's Mathematics Scaled Score is **300** which is at the **Advanced Level**. Your student's possible range of scores is from 290 to 300.

Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and are able to:

- Select and use multiple problem-solving strategies to solve complex problems involving the four operations and clearly communicate strategies.
- Read, identify, and interpret place value of numbers to 1 million and apply in problem-solving situations.
- Solve multi-step addition and subtraction problems involving whole numbers and decimals with multiple regroupings.
- Solve multi-step multiplication problems with multi-digit numbers and multiple regroupings.
- Divide by two-digit divisor and interpret remainder.
- Solve problems involving addition and subtraction of simple fractions with common denominators and recognize equivalent fractions.
- Use and apply strategies and procedures to solve complex multi-step algebraic problems involving equations, number patterns, geometric patterns, and change and clearly describe the relationship.
- Use properties and extensive vocabulary to describe and identify two- and three-dimensional figures and the relationships among them.
- Solve complex geometric problems involving points on coordinate grids, symmetry, transformations, and visual and spatial reasoning and clearly communicate strategies.
- Apply tools, procedures, and formulas of measurement to solve complex problems.
- Collect, organize, display, read, and interpret complex data; use data in problem-solving situations; and find all possible outcomes of a problem.

Your student's Reading Scaled Score is **300** which is at the **Advanced Level**. Your student's possible range of scores is from 290 to 300.

Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems. Using grade-level text or above, the student is able to:

- Use advanced reading vocabulary.
- Understand main idea and support with details.
- Use prior knowledge to make meaning of text.
- Read a variety of materials.
- Understand personification, figurative language, and literary devices.
- Distinguish fact from opinion; and make inferences.
- Identify author's purpose.
- Read and interpret maps, charts and diagrams.
- Interpret and respond to text.
- Analyze, organize, and synthesize information.
- Critically evaluate text.
- Use resource materials.
- Describe reading successes and set reading goals.

Your student's Science Scaled Score is **268** which is at the **Proficient Level**. Your student's possible range of scores is from 260 to 276.

Students at this level demonstrate a solid understanding of challenging subject matter and are able to:

- With direction, safely complete a simple investigation by asking questions using identified variables, use appropriate tools, communicate results, and identify that observation is a key inquiry process used by Montana American Indians.
- Select and use tools for simple measurement of solids, liquids, and gases, identifying properties of each state of matter and describing and modeling characteristics of and changes within basic physical and mechanical systems.
- Identify attributes of biotic (living) and abiotic (non-living) objects, including classification based on similarities and differences, basic structure and function, and processes of each system.
- Identify and accurately illustrate Earth's features, locating several observable changes of those features.
- Identify interactions among technology, science, and society.
- Discuss scientific information related to current events and local problems.
- Identify the historical significance of scientists, identify the impact of their discoveries on humans today, and identify influences of science and technology on the development of Montana American Indian cultures.
- Identify examples of Montana American Indian contributions to scientific and technological knowledge.

Scores on Montana Content Standards

CRT results are reported for Montana Content Standards in Mathematics, Reading, and Science to provide standard-specific information about the student's achievement. The results can be used to show the student's relative performance on the standards within a content area.

Mathematics	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Problem Solving	This standard is assessed within the frameworks of standard 2-7.		
2. Numbers and Operations	22	15	8-21
3. Algebra	8	8	1-8
4. Geometry	10	6	0-10
5. Measurement	10	9	2-10
6. Data Analysis, Statistics, and Probability	8	8	1-8
7. Patterns, Relations, and Functions	8	8	0-8
Reading	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Students construct meaning as they comprehend, interpret, and respond to what they read.	21	17	6-19
2. Students apply a range of skills and strategies to read.	19	16	5-17
3. Students set goals, monitor, and evaluate their reading progress.	This standard is not measurable in a statewide assessment.		
4. Students select, read, and respond to print and nonprint material for a variety of purposes.	10	9	2-10
5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences.	10	8	0-10
Science	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Scientific Investigations	14	8	4-14
2. Physical Science	14	7	3-14
3. Life Science	14	12	5-14
4. Earth/Space Science	14	13	4-14
5. Impact on Society	Subscores are not reported for this standard.		
6. Historical Development	Subscores are not reported for this standard.		

MontCAS CRT

School: Demonstration School 1
System: Demonstration District A
Grade: 04
Spring 2011

Mathematics

School Summary Report

Confidential

I. Distribution of Scores

Perf. Level	Scores	School		% of Students in Cat.	System		% of Students in Cat.	State		% of Students in Cat.
		N	% of Students		N	% of Students		N	% of Students	
Advanced	299-300	7	39	44	11	29	37	2587	25	31
	297-298	0	0		0	0		0	0	
	295-296	1	6		2	5		339	3	
	293-294	0	0		1	3		382	4	
	291-292	0	0		0	0		0	0	
Proficient	283-290	0	0	28	2	5	26	738	7	39
	275-282	3	17		4	11		1030	10	
	266-274	1	6		2	5		942	9	
	258-265	1	6		2	5		590	6	
	250-257	0	0		0	0		790	7	
Nearing Proficiency	245-249	1	6	11	3	8	21	445	4	17
	240-244	0	0		0	0		451	4	
	235-239	0	0		0	0		189	2	
	230-234	1	6		4	11		360	3	
	225-229	0	0		1	3		300	3	
Novice	220-224	1	6	17	2	5	16	298	3	13
	215-219	0	0		1	3		125	1	
	210-214	0	0		0	0		202	2	
	205-209	0	0		0	0		198	2	
	200-204	2	11		3	8		579	5	

II. Subtest Results

Mathematics		Possible Points	Average Points Earned		
			School	System	State
Total Points		66	44	43	43
Standards	1. Problem Solving	This standard is assessed within the frameworks of standards 2–7.			
	2. Numbers and Operations	22	14	14	15
	3. Algebra	8	6	6	6
	4. Geometry	10	6	6	6
	5. Measurement	10	6	6	7
	6. Data Analysis, Statistics, and Probability	8	5	5	5
	7. Patterns, Relations, and Functions	8	5	5	5

CRT Performance Level Descriptors

Advanced (291-300)

This level denotes superior performance.

Proficient (250-290)

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)

This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)

This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

MontCAS CRT

Confidential

Mathematics

**School
Summary
Report**

School: Demonstration School 1
System: Demonstration District A
Grade: 04
Spring 2011

III. Results for Subgroups of Students

Reporting Category	School					System					State				
	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A
All Students	18	17	11	28	44	38	16	21	26	37	10545	13	17	39	31
Gender															
Male	8	*	*	*	*	20	10	20	15	55	5358	14	16	37	33
Female	10	20	10	40	30	18	22	22	39	17	5187	13	17	40	30
Ethnicity															
American Indian or Alaska Native	3	*	*	*	*	7	*	*	*	*	1329	33	23	33	11
Asian	0	*	*	*	*	1	*	*	*	*	103	7	9	37	48
Hispanic	2	*	*	*	*	3	*	*	*	*	404	16	22	43	19
Black or African American	1	*	*	*	*	1	*	*	*	*	151	18	23	40	19
Native Hawaiian or Other Pacific Islander	0	*	*	*	*	1	*	*	*	*	38	18	8	45	29
White	12	8	8	42	42	24	13	17	33	38	8520	10	15	39	35
Special Education	1	*	*	*	*	3	*	*	*	*	1227	41	23	24	11
Students with a 504 Plan	1	*	*	*	*	1	*	*	*	*	56	13	18	36	34
Title I (optional)	9	*	*	*	*	17	35	18	18	29	4491	20	21	37	22
Tested with Standard Accommodation	3	*	*	*	*	5	*	*	*	*	1849	32	26	32	10
Tested with Non-Standard Accommodation	1	*	*	*	*	1	*	*	*	*	41	41	10	37	12
Alternate Assessment	If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report														
Migrant	1	*	*	*	*	1	*	*	*	*	27	19	30	41	11
Gifted/Talented	2	*	*	*	*	2	*	*	*	*	570	0	2	21	78
LEP/ELL	1	*	*	*	*	2	*	*	*	*	238	45	28	23	4
Former LEP Student	0	*	*	*	*	1	*	*	*	*	274	34	24	32	10
LEP Student Enrolled for First Time in a U.S. School	1	Performance levels are not reported for 1st year LEP students													
Free/Reduced Lunch	11	27	9	36	27	19	26	26	26	21	5011	20	21	39	20

*Less than ten (10) students were assessed

MontCAS CRT

System: Demonstration District A
Grade: 04
Spring 2011

Mathematics

System Summary Report

Confidential

I. Distribution of Scores

Perf. Level	Scores	System			State		
		Number	% of Students	% of Students in Cat.	Number	% of Students	% of Students in Cat.
Advanced	299-300	11	29	37	2587	25	31
	297-298	0	0		0	0	
	295-296	2	5		339	3	
	293-294	1	3		382	4	
	291-292	0	0		0	0	
Proficient	283-290	2	5	26	738	7	39
	275-282	4	11		1030	10	
	266-274	2	5		942	9	
	258-265	2	5		590	6	
	250-257	0	0		790	7	
Nearing Proficiency	245-249	3	8	21	445	4	17
	240-244	0	0		451	4	
	235-239	0	0		189	2	
	230-234	4	11		360	3	
	225-229	1	3		300	3	
Novice	220-224	2	5	16	298	3	13
	215-219	1	3		125	1	
	210-214	0	0		202	2	
	205-209	0	0		198	2	
	200-204	3	8		579	5	

II. Subtest Results

Mathematics		Possible Points	Average Points Earned	
			System	State
Total Points		66	43	43
Standards	1. Problem Solving	This standard is assessed within the frameworks of standards 2–7.		
	2. Numbers and Operations	22	14	15
	3. Algebra	8	6	6
	4. Geometry	10	6	6
	5. Measurement	10	6	7
	6. Data Analysis, Statistics, and Probability	8	5	5
	7. Patterns, Relations, and Functions	8	5	5

CRT Performance Level Descriptors

Advanced (291-300)

This level denotes superior performance.

Proficient (250-290)

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)

This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)

This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

III. Results for Subgroups of Students

Reporting Category	System					State				
	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A
All Students	38	16	21	26	37	10545	13	17	39	31
Gender										
Male	20	10	20	15	55	5358	14	16	37	33
Female	18	22	22	39	17	5187	13	17	40	30
Ethnicity										
American Indian or Alaska Native	7	*	*	*	*	1329	33	23	33	11
Asian	1	*	*	*	*	103	7	9	37	48
Hispanic	3	*	*	*	*	404	16	22	43	19
Black or African American	1	*	*	*	*	151	18	23	40	19
Native Hawaiian or Other Pacific Islander	1	*	*	*	*	38	18	8	45	29
White	24	13	17	33	38	8520	10	15	39	35
Special Education	3	*	*	*	*	1227	41	23	24	11
Students with a 504 Plan	1	*	*	*	*	56	13	18	36	34
Title I (optional)	17	35	18	18	29	4491	20	21	37	22
Tested with Standard Accommodation	5	*	*	*	*	1849	32	26	32	10
Tested with Non-Standard Accommodation	1	*	*	*	*	41	41	10	37	12
Alternate Assessment	If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report									
Migrant	1	*	*	*	*	27	19	30	41	11
Gifted/Talented	2	*	*	*	*	570	0	2	21	78
LEP/ELL	2	*	*	*	*	238	45	28	23	4
Former LEP Student	1	*	*	*	*	274	34	24	32	10
LEP Student Enrolled for First Time in a U.S. School	1	Performance levels are not reported for 1st year LEP students								
Free/Reduced Lunch	19	26	26	26	21	5011	20	21	39	20

*Less than ten (10) students were assessed

MontCAS CRT

Grade: 04
Spring 2011

Mathematics

State Summary Report

Confidential

I. Distribution of Scores

Perf. Level	Scores	State		
		Number	% of Students	% of Students in Cat.
Advanced	299-300	2587	25	31
	297-298	0	0	
	295-296	339	3	
	293-294	382	4	
	291-292	0	0	
Proficient	283-290	738	7	39
	275-282	1030	10	
	266-274	942	9	
	258-265	590	6	
	250-257	790	7	
Nearing Proficiency	245-249	445	4	17
	240-244	451	4	
	235-239	189	2	
	230-234	360	3	
	225-229	300	3	
Novice	220-224	298	3	13
	215-219	125	1	
	210-214	202	2	
	205-209	198	2	
	200-204	579	5	

II. Subtest Results

Mathematics		Possible Points	Average Points Earned
			State
Total Points		66	43
Standards	1. Problem Solving	This standard is assessed within the frameworks of standards 2–7.	
	2. Numbers and Operations	22	15
	3. Algebra	8	6
	4. Geometry	10	6
	5. Measurement	10	7
	6. Data Analysis, Statistics, and Probability	8	5
	7. Patterns, Relations, and Functions	8	5

CRT Performance Level Descriptors

Advanced (291-300)

This level denotes superior performance.

Proficient (250-290)

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)

This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)

This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

MontCAS CRT

Confidential

Mathematics

**State
Summary
Report**

Grade: 04
Spring 2011

III. Results for Subgroups of Students

Reporting Category	State				
	Number	% in N	% in NP	% in P	% in A
All Students	10545	13	17	39	31
Gender					
Male	5358	14	16	37	33
Female	5187	13	17	40	30
Ethnicity					
American Indian or Alaska Native	1329	33	23	33	11
Asian	103	7	9	37	48
Hispanic	404	16	22	43	19
Black or African American	151	18	23	40	19
Native Hawaiian or Other Pacific Islander	38	18	8	45	29
White	8520	10	15	39	35
Special Education	1227	41	23	24	11
Students with a 504 Plan	56	13	18	36	34
Title I (optional)	4491	20	21	37	22
Tested with Standard Accommodation	1849	32	26	32	10
Tested with Non-Standard Accommodation	41	41	10	37	12
Alternate Assessment		*	*	*	*
Migrant	27	19	30	41	11
Gifted/Talented	570	0	2	21	78
LEP/ELL	238	45	28	23	4
Former LEP Student	274	34	24	32	10
LEP Student Enrolled for First Time in a U.S. School	6	Performance levels are not reported for 1st year LEP students			
Free/Reduced Lunch	5011	20	21	39	20

*Less than ten (10) students were assessed

Appendix R—SAMPLE INTERACTIVE REPORTS



Confidential

Roster and Item-Level Report

Mathematics

System: Demonstration District A
 School: Demonstration School 1
 Grade: 04
 Date: 8/25/2011

Page: 1 of 1

Released Item Number Content Standard Depth of Knowledge Code Item Type Correct MC Response		Released Items																														Total Test Results																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Points Earned by Standard on CRT						Total Points Earned on CRT	Scaled Score	Performance Level									
		2	6	7	5	2	2	3	2	7	2	2	6	3	6	5	4	7	2	2	4	5	6	5	3	2	2	6	7	3	4	Standard 2	Standard 3	Standard 4	Standard 5	Standard 6	Standard 7												
		1	2	2	2	2	2	2	1	2	1	2	2	1	2	1	2	2	2	2	2	2	2	2	2	1	2	2	1	2	MC										MC	MC	MC	MC	MC	MC	MC	MC	CR
		A	D	D	C	D	C	D	C		A	B	D	B	C	D	A	B	C		D	D	A	C	B	C	A	A	C	A																			
Name/Student ID	Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4		22	8	10	10	8	8	66											
CHANG, BRYANT	D04100020	+	B	+	A	+	+	+	+	1	+	+	+	+	+	C	+	+	+	1	+	+	+	+	+	+	+	+	+	+	+	+	+	3	20	7	9	7	6	7	56	300	A						
CHILCOTT, LINDA	D04100003	+	+	+	+	+	+	+	D	1	+	+	+	+	+	+	C	+	+	0	+	+	+	+	+	+	+	+	+	+	+	+	+	3	15	8	6	9	8	8	54	300	A						
DAVIS, MADISON	D04100012									B										B	B	B	B	+	A	+	D	B	B	+	B		1	1	1	1	0	1	5	200	N								
DEMELLO, SARAH	D04100035	+	+	C	+	C	B	+	D	0	+	C	+	D	B	B	+	+	A	0	A	B	D	+	A	B	C	C	+	+	0	3	5	3	3	4	3	21	203	N									
FISHER, DERRICK	D04100001	+	C	C	+	C	+	+	+	0	+	+	+	C	+	+	D	+	+	1	+	+	D	+	+	+	C	B	B	B	3	18	6	6	10	5	5	50	289	P									
FONTESSILIE, DESTINY	D04100008	+	+	+	D	B	D	+	+	1	+	+	+	+	+	C	+	+	+	1	C	A	C	A	+	+	C	D	+	+	3	14	7	6	5	5	5	42	263	P									
GARCIA, MATTHEW	D04100002									B										B														B					0		ALT								
GARGAN, SHAWNA	D04100006	+	+	+	+	+	+	C	+	1	+	C	+	C	+	+	+	+	+	1	+	+	+	A	+	+	+	B	+	+	4	20	6	9	9	6	8	58	300	A									
HALLIDAY, DEVON	D04100030	+	+	+	B	+	+	+	+	0	+	+	C	+	+	+	+	+	D	1	+	+	+	+	+	+	B	+	+	+	3	18	8	8	7	6	7	54	300	A									
HANGE, SHAUN	D04100005	+	+	+	+	C	+	+	+	1	+	+	+	+	+	+	+	+	+	1	+	+	+	B	+	+	+	+	+	+	3	19	7	8	8	7	8	57	300	A									
HOLLAND, ADDIE	D04100027	+	A	B	+	A	+	+	+	0	+	C	+	D	D	C	D	+	A	1	A	+	B	A	+	+	B	B	+	+	2	8	4	3	4	2	2	23	208	LEP									
HORNE, DOUGLAS	D04100028	+	+	+	+	+	+	+	+	1	+	+	+	+	+	C	+	+	+	1	+	+	+	+	+	+	+	+	+	+	3	22	8	8	7	7	8	60	300	A									
JARRETT, JESSICA	D04100049	+	+	C	+	+	B	C	D	0	+	C	+	+	+	C		A	+	1	+	+	D	+	+	+	C	D	A	+		10	6		8	5	2	33	243	NP									
JOHNSON, NICHOLAS	D04100017	B	A	B	A	B	B	B	+	0	+	A	+	+	+	C	+	+	+	0	+		+	B	C	+	+	C	A	B	2	10	5	6	1	3	3	28	223	N									
LIBERTY, AMANDA	D04100047	+	+	+	+	+	D	+	+	0	+	C	C	A	+	C	D	C	+	0	+	+	+	+	+	+	+	+	+	+	3	10	6	6	7	5	3	37	249	NP									
LINGLE, JEFFREY	D04100024	+	+	+	+	+	+	+	+	1	+	C	C	+	+	C	B	+	+	1	+	+	+	D	+	+	+	D	+	+	3	19	8	7	6	5	8	53	299	A									
MILLER, TYLER	D04100009	B	+	+	+	+	+	+	+	0	+	+	+	+	+	C	D	D	+	1	+	+	+	A	+	+	D	B	B	+	3	16	6	7	8	7	3	47	279	P									
MONTOYA, ALCY		C	C	+	+	+	+	+	+	0	+	A	C	C	+	C	+	D	A	1	+	C	+	D	+	+	+	B	+	+	0	10	6	4	5	3	3	31	231	NP									
THOMPSON, JOSHUA	D04100031	+	+	+	+	+	D	+	D	1	D	+	B	+	D	+	D	+	B	1	+	+	+	+	+	+	C	+	+	+	2	11	6	6	9	5	7	44	270	P									
TIPPING, EMILY	D04100045	+	+	+	+	+	A	+	+	1	+	C	+	C	+	C	+	D	+	1	+	+	C	+	+	+	+	B	+	+	3	15	7	8	8	6	4	48	282	P									
WIGLEY, CAITLIN	D04100013	+	+	+	+	+	+	+	+	1	+	+	+	+	+	C	+	A	+	1	+	C	D	A	+	+	+	+	+	+	4	16	7	7	5	6	5	46	276	P									
WILLISON, TREVOR	D04100046	+	+	+	+	+	+	+	+	0	+	+	+	+	+	C	+	+	+	1	C	+	+	+	+	+	C	+	+	+	3	18	7	7	6	7	7	52	296	A									
																																							</										



Performance Level Summary

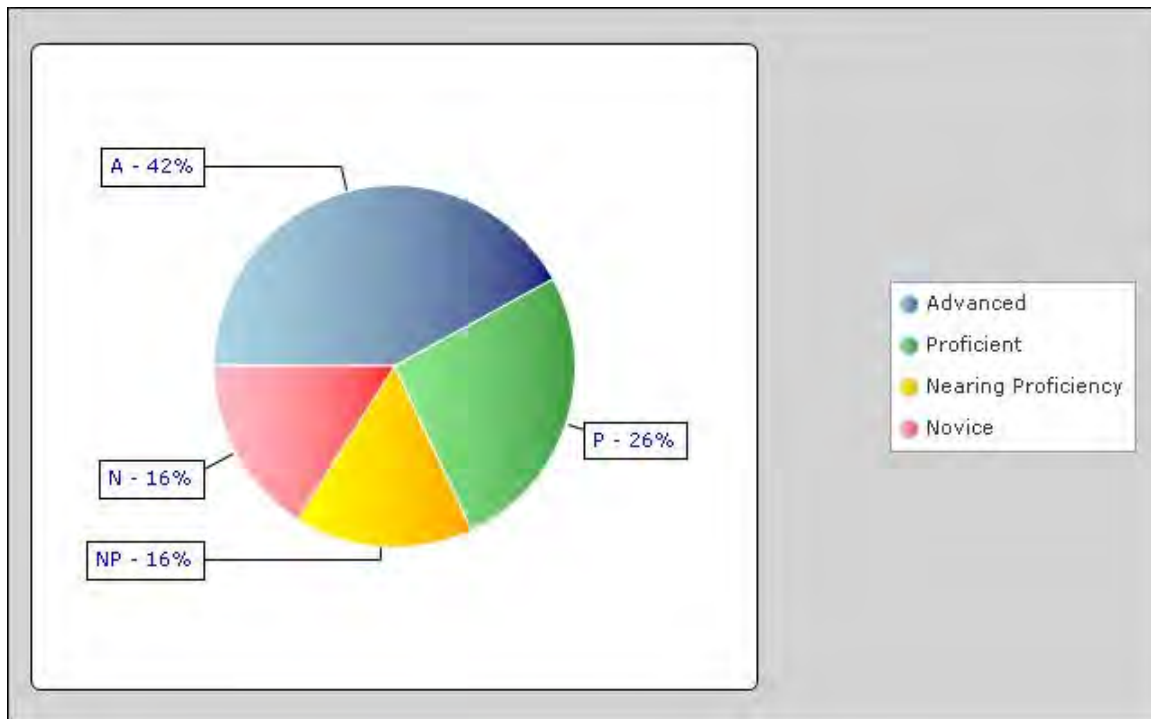
System: Demonstration District A

School: Demonstration School 1

Grade: 04

Date: 8/25/2011 1:44:07 PM

Mathematics



Performance Level	Count	Percentage %*
Advanced	8	42
Proficient	5	26
Nearing Proficiency	3	16
Novice	3	16

*Percentages may not total exactly 100% due to applied rounding.



C O N F I D E N T I A L

Student Name
BRYANT CHANG

Longitudinal Data Report

Year	Enrolled Grade	School Name	Administration	Test Name	Content Area	Score	Achievement Level
0910	04	Demonstration School 2	MontCAS CRT	Grade 04 Mathematics	mat	292	Advanced
0910	04	Demonstration School 2	MontCAS CRT	Grade 04 Reading	rea	300	Advanced
0910	04	Demonstration School 2	MontCAS CRT	Grade 04 Science	sci	279	Proficient
1011	04	Demonstration School 1	MontCAS CRT	Grade 04 Mathematics	mat	300	Advanced
1011	04	Demonstration School 1	MontCAS CRT	Grade 04 Reading	rea	300	Advanced
1011	04	Demonstration School 1	MontCAS CRT	Grade 04 Science	sci	292	Advanced

Note: This report returns as many years of data as are available for this student beginning with 06-07.



Mathematics Item Analysis Summary

System: Demonstration District A

School: Demonstration School 1

Grade: 04

Date: 8/25/2011 1:44:46 PM

Multiple Choice

Released Item	Standard	Correct (#)	A (#)	B (#)	C (#)	D (#)	IR (#)	Correct Response
1	2	14	14	2	1	0	1	A
2	6	14	1	1	1	14	1	D
3	7	15	0	1	1	15	1	D
4	5	13	2	1	13	1	1	C
5	2	13	0	2	2	13	1	D
6	2	11	1	2	11	3	1	C
7	3	15	0	1	1	15	1	D
8	2	14	0	0	14	3	1	C
10	2	16	16	0	0	1	1	A
11	2	10	2	10	5	0	1	B
12	6	12	0	1	4	12	1	D
13	3	12	1	12	3	1	1	B
14	6	15	0	1	15	1	1	C
15	5	5	0	1	11	5	1	D
16	4	12	12	1	1	3	1	A
17	7	12	1	12	1	3	1	B
18	2	13	2	1	13	1	1	C
20	4	14	1	1	2	14	0	D
21	5	12	1	2	2	12	1	D
22	6	13	13	1	2	2	0	A
23	5	10	4	2	10	2	0	C
24	3	15	2	15	1	0	0	B
25	2	17	0	1	17	0	0	C
26	2	11	11	1	4	2	0	A
27	6	9	9	5	2	2	0	A
28	7	15	1	2	15	0	0	C
29	3	17	17	1	0	0	0	A

Constructed Response

Released Item	Standard	Point Value	Average Score
9	7	1	0.6
19	2	1	0.7
30	4	4	2.5